AMERICAN

JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LV
1933

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

HYPERELLIPTIC FUNCTIONS AND IRRATIONAL BINARY INVARIANTS. II.

By Arthur B. Coble.

The first article [8] of this series dealt primarily with the linear dependence of the invariants (A), and the invariants (B), of the binary $(2p+2)$-ic, and with the modular manifolds, $M_{2p-1}(x)$, $M_{2p-1}(\xi)$, defined respectively by them. In the present article the particular case, $p=3$, is discussed in some detail with an occasional generalization. Sections and references in the two are numbered consecutively, only those references being repeated which are cited anew.

The case $p=3$ shares with the case $p=2$ the peculiarity that $g_{(2p+2)!}$ contains subgroups of low index, for $g_{8!}$ the index being 30, and for $g_{8!/2}$ the index being 15. These subgroups can be defined by invariants which are linear in the invariants (A), and invariants linear in the invariants (B). Thus the invariants of these subgroups constitute a natural coördinate system for a discussion of the invariants (A) and (B), and the modular manifolds $M_5(x)$, $M_5(\xi)$ [cf. **9**]. The tactical relations of these invariants are developed in **7**. On the other hand, the case $p=3$ is more typical of the general case than $p=2$, since the modular manifolds cannot be expressed by a single equation.

In contrast to the case $p=2$, the hyperelliptic modular functions for $p=3$ are special functions characterized by the fact that the even theta function, $\vartheta(u)$, vanishes for the zero argument. It is desirable therefore to have, as well, a treatment [cf. **8**] parallel to that of the generic functions attached to a planar quartic curve as given, for example, in ([3], pp. 192-5).

7. Tactical configurations, $p=3$. Let the 63 half periods of the generic theta functions $(p=3)$, or the discriminant factors of the generic quartic curve, be represented by the points of a finite space modulo two, $S_5(2)$, in which a null system N is given. These points can be named in a basis notation, P_{ij}, $P_{ijkl}=P_{mnop}$ $(i,\cdots,p=1,\cdots,8)$ [cf. [3], §§ 22, 24, 25; in particular, p. 68]. The 35 points P_{ijkl} are on a quadric Q associated with $\vartheta(u)$ whose polar system is N, and the 28 points P_{ij} are not on Q. The collineation group in $S_5(2)$ which leaves N unaltered, the modular group for generic $p=3$, has the order $8!\,36$. The subgroup which leaves Q un-

altered, the modular group for hyperelliptic $p=3$, has the order $8!$. It effects on the points the permutation group $g_{8!}$ of their subscripts, $i, \cdots, p$.

The null system N has 315 null lines which divide into 210 of type P_{ij}, P_{kl}, P_{ijkl} tangent to Q, and 105 of type P_{ijkl}, P_{ijmn}, P_{klmn} contained in Q. The null system N has 135 null planes, or Göpel planes, which divide into 105 of type

$$\text{(a)} \qquad P_{12},\, P_{34},\, P_{56},\, P_{78},\, P_{1234},\, P_{1256},\, P_{1278},$$

which touch Q along a generator; and 30 of type

$$\text{(b)} \qquad P_{1234},\, P_{1256},\, P_{1278},\, P_{1468},\, P_{2368},\, P_{2458},\, P_{2467},$$

which are contained in Q.

The outstanding facts concerning the geometry on the quadric Q in $S_5(2)$ are immediate consequences of the fact that Q is the map of degenerate null systems, or lines, in the finite space $S_3(2)$. For, the generic null system in $S_3(2)$, $\Sigma a_{ik}(x_i y_k - x_k y_i) = 0$ $(i, k = 1, \cdots, 4;\ i < k)$, is degenerate, i. e., is a line in $S_3(2)$, if $a_{12}a_{34} + a_{13}a_{42} + a_{14}a_{23} \equiv 0$ (mod. 2). But this latter congruence is precisely an equation of Q in terms of the six coördinates a_{ik} in $S_5(2)$. Thus the 35 lines in $S_3(2)$ map into the 35 points on Q. The collineation group in $S_3(2)$ has the order $8!/2$, and it is isomorphic with the even permutations of eight things. The correlation group in $S_3(2)$ has the order $8!$, and it is isomorphic with all the permutations of eight things. This correlation group maps into the $g_{8!}$ which leaves Q unaltered.

The 15 points in $S_3(2)$, each on seven lines, map into 15 Göpel planes on Q of type (b), each on seven points. The 15 planes in $S_3(2)$, each on seven lines, map into 15 Göpel planes on Q also of type (b), each on seven points. Since a correlation in $S_3(2)$ interchanges the points and planes of $S_3(2)$, the 30 Göpel planes on Q divide into two conjugate sets of 15 each under $g_{8!/2}$. If a Göpel plane of one set of 15 be selected, and a line in it be isolated, there is a unique Göpel plane of the other set of 15 which contains the isolated line. In this way a triad of Göpel planes, one of a set of 105, appears, as e. g. in

$$\text{(c)} \qquad \begin{array}{lllllll} P_{12}, & P_{34}, & P_{56}, & P_{78}, & P_{1234}, & P_{1256}, & P_{1278}; \\ P_{1468}, & P_{2368}, & P_{2458}, & P_{1358}, & P_{1234}, & P_{1256}, & P_{1278}; \\ P_{2468}, & P_{1368}, & P_{1458}, & P_{2358}, & P_{1234}, & P_{1256}, & P_{1278}. \end{array}$$

The corresponding figure in $S_3(2)$ is the self-dual figure of three lines of a plane pencil.

To the one set of 15 Göpel planes which maps as above the 15 points of $S_3(2)$ we attach in **9** a set of 15 linearly related invariants, $\sigma_1, \cdots, \sigma_{15}$;

and to the other set of 15 Göpel planes which maps the 15 planes of $S_3(2)$ a similar set of 15 linearly related invariants, $\bar{\sigma}_1, \cdots, \bar{\sigma}_{15}$. All of the Göpel invariants can be expressed linearly in terms of the 15 σ's. Each of the thirty $\sigma, \bar{\sigma}$'s (or Göpel planes) is invariant under a $g_{8.168}$ in $g_{8!}$. Under the even subgroup $g_{8!/2}$ each set of σ's is invariant; under the odd elements of $g_{8!}$ the two sets of σ's interchange. In **10** we define in terms of the invariants (B) similar sets of irrational invariants $\tau_1, \cdots, \tau_{15}$; $\bar{\tau}_1, \cdots, \bar{\tau}_{15}$. Each of these admits a $g_{8.168}$, and thus they are in one-to-one correspondence with $\sigma, \bar{\sigma}$.

A line in $S_3(2)$ is on three points and three planes, any two of either set determining the line. Hence on a point of Q there are three Göpel planes from each set of 15. Thus in each set of 15 there is a triad system containing 35 triads, and a triad in one set is paired with a triad in the other set.

The duality in $S_3(2)$ between point and plane is reflected on Q by the fact that each Göpel plane of one set determines seven of the other set. Thus in (c) if the second Göpel plane be fixed, and the line in $S_5(2)$ common to the triad be allowed to vary, the third Göpel plane runs over a set of seven of the other set.

The notation of the next section has the following origin in $S_3(2)$. If degenerate null systems in $S_3(2)$ are mapped as above on the points of Q in $S_5(2)$, then non-degenerate null systems in $S_3(2)$, of which there are 28, are mapped upon the points of $S_5(2)$ not on Q. These are in one-to-one correspondence with the 28 discriminant factors of the underlying octavic. If, in particular, the difference, $(78) = (t_7 - t_8)$, and the null system N_{78}, be isolated, the $g_{8!/2}$ reduces to the $g_{6!}$ associated with the case $p = 2$. The invariants (A) which contain the factor (78) can all be expressed linearly in terms of six linearly related invariants, $a, \cdots, f$ (cf. [3], p. 114).

These tactical relations are discussed in great detail by M. Noether [9] and E. H. Moore.[10] We have developed here only so much as will be useful later.

8. The 135 hyperelliptic Göpel invariants and the linear relations which connect them. In the generic case, $p = 3$, the product of the seven discriminant factors of the underlying ternary quartic curve, which are attached to the seven points of a Göpel plane, is a Göpel invariant. The 135 Göpel invariants satisfy a system of 315 three-term relations (corresponding to the 315 null lines in $S_5(2)$) by means of which they may be expressed with numerical coefficients in terms of 15 which are linearly independent. In the hyperelliptic case, however, there are but 28 discriminant factors of the underlying binary octavic, namely, $(ij) = (t_i - t_j)$. These are attached to the 28 points P_{ij} in $S_5(2)$ not on Q. We proceed to construct for this

special case a new set of Göpel invariants which have linear properties precisely like those which obtain in the general case [cf. [3], pp. 192-5].

To the 105 Göpel planes of the type **7** (a) it is evidently suitable to attach Göpel invariants of the type (12)(34)(56)(78). For, if three Göpel planes of this type meet in a null line, such as P_{12}, P_{34}, P_{1234}, the corresponding Göpel invariants satisfy a three-term relation. Of these 105 invariants, 15 contain the factor (78). These are expressible in terms of five, or of six which are linearly related. We denote such a set of six by $a, \cdots, f$, and set

$$[ab] = (a+b) = 5(15)(24)(36)(78), \tag{1}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

This is a sample of 15 Göpel invariants which arise from (1) by applying the parallel permutations:

$$\begin{aligned}(12) &: (ad)(be)(cf),\\ (23456) &: (adbfe),\end{aligned} \tag{2}$$

an odd permutation being accompanied by a change of sign in $a, \cdots, f$.

By examining $a+b$, $c+d$, $e+f$, we find that

$$a+b+c+d+e+f=0, \quad \text{or} \tag{3}$$

$$[ab]+[cd]+[ef]=0. \tag{α}$$

It is also evident that

$$[ab][ac][bc] = [de][df][ef], \tag{4}$$

since each term involves the same discriminant factors. By using (1) and (3) this cubic relation may be converted into

$$a^3+b^3+c^3+d^3+e^3+f^3=0. \tag{5}$$

Any set of numbers, $a:b:\cdots:f$, which satisfy (3) and (5), determine an ordered binary sextic with roots projective to $t_1, \cdots, t_6$ [cf. [3], § 35].

With the seven lines of the second and third Göpel planes in **7** (c) as a guide we form the two linear invariants,

$$\begin{aligned}-[ad-] = {} & (12)(34)(56)(78)+(14)(32)(57)(68)\\ & +(16)(52)(37)(48)+(36)(54)(17)(28)\\ & +(13)(42)(67)(58)+(15)(62)(47)(38)\\ & +(35)(64)(27)(18),\end{aligned}$$

$$\begin{aligned}-[ad+] = {} & (12)(34)(56)(78)+(24)(13)(57)(68)\\ & +(26)(15)(37)(48)+(36)(45)(27)(18)\\ & +(23)(14)(67)(58)+(25)(16)(47)(38)\\ & +(35)(46)(17)(28).\end{aligned} \tag{6}$$

The second of these two arises from the first by applying the transposition (78), and by changing the sign. Each admits the same $g_{8.168}$ as the corresponding Göpel plane. In each every difference (ij) occurs just once. In particular, the difference (78) occurs with the same group, $(12)(34)(56)$, as in $[ad]$ in (1). By operating on (6) with the group (2), thirty new Göpel invariants are obtained to complete the set of 135.

In the identity,

$$(12)(34)\cdot(56)(78) = [(13)(24) - (14)(23)]\,[(57)(68) - (58)(67)],$$

the four terms on the right are found in (6). This, and the similar identities formed for $(12)(56)\cdot(34)(78)$ and $(12)(78)\cdot(34)(56)$, yield the three-term relation,

$$(\beta) \qquad [ad+] + [ad-] + [ad] = 0.$$

In $[ad+]$, $[ad-]$, $[ad]$ we have 45 of the 135 Göpel invariants including 15 of the type $(12)(34)(56)(78)$. The remaining 90, all of this latter type, are defined by

$$(\gamma) \qquad [ad, be] + [ad-] + [be+] = 0.$$

Again the members of this set of 90 are obtained from any one (e. g., the one in (8) below) by the operations of the group (2). From the definition of $[ad, be]$ in (γ) we find, by using (α) and (β), that

$$(\delta) \qquad [ad, be] + [be, ad] + [cf] = 0,$$

$$(\epsilon) \qquad [ad, be] + [be, cf] + [cf, ad] = 0.$$

The relations $(\alpha), \cdots, (\epsilon)$ include 15, 15, 90, 45, 30 of the 315 three-term relations which correspond in $S_5(2)$ to the sets of three Göpel planes on the 315 null lines. We have yet to prove that the remaining 120 relations have the form:

$$(\zeta) \qquad [ab, de] + [bc, ef] + [ca, fd] = 0.$$

For this purpose an explicit expression for the type $[ab, de]$ defined by (γ) is necessary. Writing (β) in the form,

$$(7) \qquad [ad] = -[ad+] - [ad-] = 5(12)(34)(56)(78),$$

we effect the permutation $(27)(35)$ which leaves $[ad-]$ unaltered. It carries $[ad+]$ into $[x, y+]$, where the letters, x, y are to be determined from the term in $[ad+]$ which contains the factor (78). This term arises from the term in $[ad+]$ which contains the factor (28), and this term therefore is $(53)(46)(12)(78)$ which occurs in $[be]$ in (1). Thus $[xy+]$

is $[be+]$. On applying the permutation to the right member of (7), we find, by using (γ), that

$$(8) \qquad [ad, be] = -[ad-] - [be+] = 5(17)(54)(36)(28).$$

From (7) and (8) we conclude that

(9) *The Göpel invariants,* $[ij+]$, $[kl-]$, *have a term* T *in common, if their literal indices* ij, kl, *have an even number* (0 *or* 2) *of letters in common. Their sum is then* $5T$.

Thus the terms in (ζ), obtained by applying the group (2) to the formula (8), are

$$(10) \qquad \begin{aligned} [de, ab] &= -[de-] - [ab+] = 5(12)(73)(45)(68), \\ [ef, bc] &= -[ef-] - [bc+] = 5(12)(73)(56)(48), \\ [fd, ca] &= -[fd-] - [ca+] = 5(12)(73)(64)(58), \end{aligned}$$

which proves the validity of (ζ).

If the theorem (9) is applied to the seven terms in $[ab+]$, and the results added, we obtain

$$(11) \qquad \begin{aligned} 2[ab+] + [ab-] + [cd-] + [ce-] & \\ + [cf-] + [de-] + [df-] + [ef-] &= 0. \end{aligned}$$

Similarly

$$(12) \qquad \begin{aligned} 2[ab-] + [ab+] + [cd+] + [ce+] & \\ + [cf+] + [de+] + [df+] + [ef+] &= 0. \end{aligned}$$

On adding the 15 relations in each set, we have

$$2\Sigma_{15}[ab+] + 7\Sigma_{15}[ab-] = 0, \qquad 2\Sigma_{15}[ab-] + 7\Sigma_{15}[ab+] = 0.$$

From this there follows

$$(13) \qquad \Sigma_{15}[ab+] = 0, \qquad \Sigma_{15}[ab-] = 0.$$

The 15 Göpel invariants, $[ab+]$, subject to the single relation (13), suffice for the linear expression of all of the invariants. For, the invariants $[ab-]$ can be expressed in terms of them as in (12), and the 105 products (A) can be expressed in terms of the two sets by means of (7), (8), (β), (γ). A similar statement applies to the 15 invariants $[ab-]$ subject to the single relation (13). On comparing the relations (α), $\cdots$, (ζ) with the like system of relations [cf. [3], p. 194 (11)] which obtains in the generic case, we see that

(14) *In the hyperelliptic case, $p=3$, the* 135 *Göpel invariants defined above satisfy the same system of* 315 *three-term relations as obtains in the generic case. They satisfy also a further system* (11), (12), (13) *by means of which the dimension of their linear system is reduced from* 15 *to* 14.

Our present notation is based on a symmetry in the letters $a, \cdots, f$ due to a symmetry in the roots $t_1, \cdots, t_6$. There is however an underlying symmetry in the roots $t_1, \cdots, t_6, t_7, t_8$. Thus formulae of the same character with respect to $g_{8!}$ take several forms in this notation, e. g., (7), (8). In order to make such transitions readily we give the effect of two generators, which extend $g_{6!}$ ot $g_{8!}$, upon the Göpel invariants.

(15) *Under the even permutation,* (27)(35), *the following pairs of Göpel invariants are interchanged:*

$$\begin{gathered}[ad+],[be+];\ [bc+],[bf+];\ [de+],[cd+];\ [ab+],[df+];\\ [af+],[ae+];\ [ac+],[bd+];\ [cf-],[be-];\ [bc-],[cd-];\\ [ef-],[ac-];\ [bf-],[af-];\ [ce-],[df-];\ [de-],[ae-].\end{gathered}$$

Under (12)(34)(56)(78) *the permutation is that effected by interchanging a and d.*

The 15 Göpel invariants $[ad-]$, and the 15 invariants $[ad+]$, correspond in the mapping described in to the 15 points and 15 planes of $S_3(2)$. The duality there mentioned in which each of one set is *on* seven of the other set is that described in (9), or that embodied more precisely in the formulae (11), (12).

The triad system in either set of 15 which corresponds to the lines in $S_3(2)$ is embodied in the two formulae,

$$(9_1)\quad \begin{aligned}&[ab+]+[cd+]+[ef+]+[ab-]+[cd-]+[ef-]=0,\\ &[ab+]+[bc+]+[ca+]+[de-]+[ef-]+[fd-]=0.\end{aligned}$$

The first exemplifies 15 pairs of triads; the second, 20. The first relation is a consequence of (α) and (β); the second, of (γ) and (ζ). These triads appear in the quadratic relations of the next section [cf. **9** (6)].

9. Quadric relations which define the modular manifolds $M_5(x)$ and $M_{2p-1}(x)$. In this section we shall sometimes refer for the sake of brevity to the 15 invariants $[ad-]$as $\sigma_1, \cdots, \sigma_{15}$, and to the 15 invariants $[ad+]$ as $\bar{\sigma}_1, \cdots, \bar{\sigma}_{15}$, where

$$(1)\qquad \sum_{i=1}^{i=15} \sigma_i = 0, \qquad \sum_{i=1}^{i=15} \bar{\sigma}_i = 0.$$

The subscript notation, $1, \cdots, 15$, has no relation to the subscript notation,

$1,\cdots,8$, of the roots of the underlying octavic. The $a,\cdots,f$ notation of the preceding section which is symmetrically related to the roots $t_1,\cdots,t_6$ through the group **8** (2) is the best that can be devised in this respect. As pointed out in **7** this amounts to the choice of one of the 28 proper null systems N_{78} in $S_3(2)$. This choice leads to alternative forms for similar relations as in **8** (16). With respect to N_{78} the $a,\cdots,f$ behave like the indices of a self-dual basis ($p=2$). Thus the 15 points of $S_3(2)$ may be denoted by $p_{ab}=p_{cdef}$, and the collinear conditions are

$$p_{ab}+p_{ac}+p_{bc}\equiv 0, \qquad p_{ab}+p_{cd}+p_{ef}\equiv 0.$$

We examine the non-linear relations which are satisfied by the 135 Göpel invariants. In the non-hyperelliptic case these are 63 cubic relations each associated with a discriminant factor [cf. [3], pp. 193-5]. In the present case, those invariants which contain one of the 28 proper discriminant factors, such as (78), can be expressed in terms of six which satisfy the linear and cubic relations, **8** (3), (5). If the cubic relation is written as in **8** (4), or also as in

$$\frac{a+b}{d+e}\cdot\frac{a+c}{d+f}\cdot\frac{b+c}{e+f}=-1,$$

then, since $(a+b)/(d+e)=-(15)(24)/(14)(25)=-D(12,54)$, etc., it appears as a consequence of the identity,

$$(2)\qquad D(12,54)\cdot D(12,46)\cdot D(12,65)=1,$$

connecting the double ratios of five points.

Consider the following pair of triads of the first type in **8** (16):

$$(3)\qquad \begin{aligned} \sigma_1&=[ad-], & \sigma_2&=[cf-], & \sigma_3&=[be-];\\ \bar\sigma_1&=[ad+], & \bar\sigma_2&=[cf+], & \bar\sigma_3&=[be+]. \end{aligned}$$

These two triads are connected by the relation,

$$(4)\qquad (\sigma_1+\sigma_2+\sigma_3)=-(\bar\sigma_1+\bar\sigma_2+\bar\sigma_3).$$

From **8** (7), (8) there follows that

$$(5)\qquad \begin{aligned} &\sigma_1+\bar\sigma_1=-5(12)(78)\cdot(34)(56), \quad \sigma_2+\bar\sigma_1=-5(28)(17)\cdot(35)(46),\\ &\qquad\sigma_3+\bar\sigma_1=-5(27)(18)\cdot(63)(54),\\ &\sigma_1+\bar\sigma_2=-5(18)(27)\cdot(53)(46), \quad \sigma_2+\bar\sigma_2=-5(21)(78)\cdot(54)(36),\\ &\qquad\sigma_3+\bar\sigma_2=-5(28)(17)\cdot(56)(43),\\ &\sigma_1+\bar\sigma_3=-5(17)(28)\cdot(36)(54), \quad \sigma_2+\bar\sigma_3=-5(27)(18)\cdot(56)(34),\\ &\qquad\sigma_3+\bar\sigma_3=-5(21)(78)\cdot(64)(53). \end{aligned}$$

This yields $(\sigma_3+\bar{\sigma}_2)/(\sigma_2+\bar{\sigma}_3)=(\sigma_2+\bar{\sigma}_1)/(\sigma_1+\bar{\sigma}_2)$, which simplifies, due to the linear relation (4), into

$$(6)\qquad \sigma_1\sigma_2+\sigma_1\sigma_3+\sigma_2\sigma_3=\bar{\sigma}_1\bar{\sigma}_2+\bar{\sigma}_1\bar{\sigma}_3+\bar{\sigma}_2\bar{\sigma}_3.$$

This quadratic relation among the Göpel invariants is symmetrically associated with the pair of triads given in (3). There is one such relation for each of the 35 points on Q in the finite $S_5(2)$, or for each of the 35 lines in the finite $S_3(2)$. We now prove that

(7) *The 63 cubic relations satisfied by the Göpel invariants in the non-hyperelliptic case become 28 cubic and 35 quadratic relations in the hyperelliptic case. A set of 135 constants which satisfy the linear relations of* **8** (14), *and these quadratic and cubic relations, is necessarily a set of Göpel invariants defined by an ordered binary octavic.*

For, let $R_{78}=0$ denote the cubic relation defined by the discriminant factor (78). Then, due to the satisfaction of $R_{78}=0$, the Göpel invariants define an ordered sextic with roots $t_1,\cdots,t_6$. Due also to $R_{68}=0$, they define an ordered sextic with roots $t_1, t_2, t_3, t'_4, t'_5, t'_7$. We have to show that $t'_4=t_4$, or that $D(12,34)=D(12,34')$. This requires that

$$(8)\qquad \frac{(13)(24)\cdot(56)(78)}{(14)(23)\cdot(56)(78)}=\frac{(13)(24')\cdot(5'7')(6'8')}{(14')(23)\cdot(5'7')(6'8')},$$

or that a quadratic relation of the form (6) exists. Thus the cubic relations ensure the existence of 28 ordered sextics, and the quadratic relations ensure that these sextics are reunited in an ordered binary octavic.

We wish to show further that the cubic relations are consequences of the quadratic relations, as well as to investigate more closely the linear dependence of the system of quadratic relations. If the $\bar{\sigma}$'s are eliminated from (6) by using **8** (11), the relation takes the form,

$$(9)\qquad R_{ad,cf,be}\equiv\sigma_1^2+\sigma_2^2+\sigma_3^2-2\sigma_2\sigma_3-2\sigma_3\sigma_1-2\sigma_1\sigma_2+r_2r_3+r_3r_1+r_1r_2=0,$$

where

$$(10)\qquad \begin{aligned} r_1&=[bc—]+[bf—]+[ce—]+[ef—],\\ r_2&=[ab—]+[ae—]+[bd—]+[de—],\\ r_3&=[ac—]+[af—]+[cd—]+[df—],\\ &\{r_1+r_2+r_3=-(\sigma_1+\sigma_2+\sigma_3)\}.\end{aligned}$$

Matching the 15 $[ad—]$'s with the 15 points of $S_3(2)$, this relation matches with one of the 35 lines in $S_3(2)$, or, according to (5), with one of the 35

invariants (B), the determinant product (1278)(3456). We observe first that the sum,

$$R_{ad,cf,be} + R_{db,ce,fa} + R_{bf,ca,ed} + R_{fe,cd,ab} + R_{ea,cb,df} \equiv 0, \tag{11}$$

vanishes identically in the arguments $[ad—]$. For, in $S_3(2)$ the five terms correspond to five lines which fill up $S_3(2)$. The terms correspond also to the five determinant products which are in a determinant identity, as may be seen at once by applying the second generator in **8** (2). In the first term of (11) [cf. (9), (10)] the 48 product terms which arise from $r_2r_3 + r_3r_1 + r_1r_2$ are those attached in $S_3(2)$ to pairs of points on a line which does not cut the line $\sigma_1, \sigma_2, \sigma_3$. Since any line of $S_3(2)$, except one of the five, will not cut two of the five, such products occur in two terms of (11). If however the product is attached to a pair of points on one of the five lines it will occur in four terms of the sum (11) with coefficient $+1$, and in the fifth as well with a coefficient -2 as in $-2\sigma_2\sigma_3$ in $R_{ad,cf,be}$. Hence the sum becomes

$$\Sigma_{15}[ad—]^2 + 2\Sigma_{105}[ij—][kl—] = \{\Sigma_{15}[ad—]\}^2 \equiv 0.$$

Thus the left members of the 35 quadratic relations satisfy the same linear relations as the 35 determinant products and only 14 are linearly independent [cf. 1 (8), (10)].

In order to express all the quadratic relations in terms of 15, which themselves are connected by one linear relation, we add the seven relations (9) attached in $S_3(2)$ to the seven lines containing σ_1. The sum is

$$6\sigma_1^2 - 6\Sigma_7(\sigma_2\sigma_3 + \sigma_3\sigma_1 + \sigma_1\sigma_2) + \Sigma_{15}[ad—]^2 + 4\Sigma_{105}[ij—][kl—].$$

If we denote the elementary symmetric functions of the 15 $[ad—]$'s, or σ's, by s_i, and use $s_1 = 0$, this sum yields the quadratic relation,

$$R_{ad} \equiv s_2 + 6\sigma_1^2 - 3\Sigma_7\sigma_2\sigma_3 = 0. \tag{12}$$

Then

$$\Sigma_{15}R_{ad} \equiv 0 \quad \text{(in the arguments } \sigma). \tag{13}$$

For, from (12), $\Sigma_{15}R_{ad} = 15s_2 + 6\Sigma_{15}\sigma_1^2 - 3s_2 = 6s_1^2 = 0$. Furthermore the relations R_{ad} in (12) yield the relations $R_{ad,cf,be}$ in (9). In fact

$$R_{ad} + R_{cf} + R_{be} = 3R_{ad,cf,be}. \tag{14}$$

If we write R_{ad} in the form

$$5\sigma_1^2 - 3\Sigma_7\sigma_2\sigma_3 + \Sigma_{91}[ij—][kl—] \quad \{ij, kl \neq ad;\ ij \neq kl\},$$

we observe that

$$\partial/\partial\sigma_1\{5\Sigma_{15}\sigma_1{}^3 - 9\Sigma_{35}\sigma_1\sigma_2\sigma_3 + 3s_3\} = 3R_{ad}.$$

Since (13) is satisfied, R_{ad} is the polar quadric of the first reference point, $\sigma_1, \cdots, \sigma_{15} = -14, 1, \cdots, 1$, as to the cubic spread. Due to $\Sigma_{15}\sigma_1{}^3 = 3s_3$, the equation of this spread may be written as

(15) $$2\Sigma_{15}\sigma_1{}^3 - 3\Sigma_{35}\sigma_1\sigma_2\sigma_3 = 3[2s_3 - \Sigma_{35}\sigma_1\sigma_2\sigma_3] = 0.$$

Hence

(16) *The* 35 *quadrics* (9), *and the* 15 *quadrics* (12) *are members of a linear system of dimension* 13, *the polar quadric system of the cubic spread* (15) *on which the modular manifold* $M_5(x)$ *is a locus of nodes.*

We wish now to prove that the cubic relations mentioned in (7), of which **8** (5) is a sample, are consequences of the quadratic relations (12). We observe first that s_3, the combinations of the 15 $[ad -]$'s three at a time, breaks up into two conjugate sets of terms under $g_{8!/2}$ according as the corresponding points in $S_3(2)$ are on a line, or make up a triangle. Hence we write

(17) $$s_3 = s_{3L} + s_{3T}.$$

The 35 terms in s_{3L} have already occurred in (15). The remaining 420 terms of s_3 make up s_{3T}. The relation (13), multiplied by σ_1, is

$$6\sigma_1{}^3 = 3\Sigma_7\sigma_1\sigma_2\sigma_3 - \sigma_1 s_2.$$

If this is summed for the fifteen terms σ_1, it becomes

$$6\Sigma_{15}\sigma_1{}^3 = 9s_{3L} = 18s_3.$$

Hence, due to the quadratic relations,

(18) $$s_{3L} = 2s_3, \qquad s_{3T} = -s_3.$$

These relations are not identities in the σ's, but hold only on the locus defined by the quadratic relations, which, as we seek to show, is the modular locus, $M_5(x)$.

The peculiarity of our present notation with respect to $S_3(2)$ is that the proper null system, N_{78}, in $S_3(2)$ is isolated. With respect to N_{78} the two pairs of points corresponding to $[ab -], [cd -]$ and $[ab -], [ac -]$ are respectively syzygetic or azygetic, i. e., their join is a null line or an ordinary line. The invariant s_{3T} of $g_{8!/2}$ divides into three parts, each invariant under the subgroup (the $g_{6!}$ of **8** (2)) which leaves N_{78} unaltered, i. e.,

(19) $$s_{3T} = s_{3T_0} + s_{3T_1} + s_{3T_2},$$

where s_{3T_j} is the sum of the combinations of three $[ad-]$'s whose corresponding points in $S_3(2)$ form a triangle with j null lines for sides. Thus the 420 terms of s_{3T} divide into 60 of type $[ab-][ac-][ad-]$ which make up s_{3T_0}, 180 of type $[ab-][cd-][bd-]$ which make up s_{3T_1}, and 180 of type $[ab-][cd-][ce-]$ which make up s_{3T_2}.

The terms which occur in the cubic relation of **8** (5), $a^3 + \cdots + f^3 = 0$, are expressed in terms of the Göpel invariants as follows:

(20) $$\begin{aligned} -2a &= [ab-] + [ac-] + [ad-] + [ae-] + [af-] \\ &= [ab+] + [ac+] + [ad+] + [ae+] + [af+] \\ &= -\{[ab] + [ac] + [ad] + [ae] + [af]\}/2. \end{aligned}$$

For, from **8** (1), (3) there follows $4a = \sum_{z=b}^{z=f} [az]$. If **8** (12) is rewritten by using **8** (β) to read $[ab] = [ab-] + \Sigma\,[xy+]$ $(x, y \neq a, b)$, and if this is summed over $b, \cdots, f$, we get $4a = \sum_{z=b}^{z=f} [az-] + 3[x, y+]$ $(x, y \neq a)$. Using **8** (13) to modify the last sum we get $4a = \Sigma_z[az-] - 3\Sigma_z[az+]$ $(z = b, \cdots, f)$. Similarly $4a = \Sigma_z[az+] - 3\Sigma_z[az-]$, and the first two equalities (20) are apparent. The last equality (20) is then easily proved. On substituting the first value (2) into the cubic relation, it takes the form

(21) $$\begin{aligned} 2\Sigma_{15}[ab-]^3 + 3\Sigma_{60}[ab-][ac-]\{[ab-] + [ac-]\} \\ + 6\Sigma_{60}[ab-][ac-][ad-] = 0. \end{aligned}$$

In order to obtain this cubic relation from the quadratic relations, we multiply the quadratic relation R_{ad} in (12) by $[ab-]$, this being attached to one of the eight points in $S_3(2)$ azygetic to $[ad-]$. The result is

$$6[ad-]^2[ab-] - 3[ab-]^2[bd-] - 3\Sigma_3 T_1 - 3\Sigma_3 T_2 + s_2[ab-] = 0,$$

where $\Sigma_3 T_1$, $\Sigma_3 T_2$ indicate three products corresponding to triangles of the types indicated. If the 120 relations of this type, corresponding to the 15 choices of R_{ad} and the 8 choices of $[ab-]$, are summed, we get

$$3\Sigma_{60}[ab-][ac-]\{[ab-] + [ac-]\} - 6s_{3T_1} - 6s_{3T_2} = 0,$$

since the terms in s_2 vanish because of **8** (13). This is

$$3\Sigma_{60}[ab-][ac-]\{[ab-] + [ac-]\} + 6s_{3T_0} - 6s_{3T} = 0 \quad \text{[cf. (19)]}.$$

But we have already found in (18), by the use of the quadratic relations alone, that $-6s_{3T} = 6s_3 = 2\Sigma_{15}[ad-]^3$. Hence this sum is precisely the left member of (21), and

(22) *The cubic relations of theorem* (7) *are consequences of the quadratic relations* (12). *The modular manifold* $M_5(x)$ *in* S_{13} *is completely defined by these quadratic relations. It is the entire nodal locus of the cubic spread* (15).

That there are 14 linearly independent quadrics on $M_5(x)$ may be proved from the mapping of an $S_5(y)$ upon $M_5(x)$ by cubic spreads with nodes at the seven points P_7^5 of a base in $S_5(y)$ [cf. **2** (7)]. For, the 105 linearly independent quadrics in S_{13} give rise in $S_5(y)$ to sextic spreads with four-fold points at P_7^5. The 7.56 conditions imposed on such a sextic, $(\alpha y)^6 = 0$, at P_7^5 must be reduced by 21. For, if a, b are any two points of P_7^5, $(\alpha a)^3(\alpha y)^3 \equiv 0$ and $(\alpha b)^3(\alpha y)^3 \equiv 0$; whence the condition, $(\alpha a)^3(\alpha b)^3 = 0$, is counted twice. Hence the number of linearly independent sextics is $462 - 7.56 + 21 = 91$, and $105 - 91 = 41$ of the quadrics in S_{13} must contain $M_5(x)$.

Similarly, the 560 cubic spreads in S_{13} give rise to spreads, $(\alpha y)^9 = 0$, in $S_5(y)$ with six-fold points at P_7^5. Again, if $(\alpha a)^4(\alpha y)^5 \equiv 0$ and $(\alpha b)^4(\alpha y)^5 \equiv 0$, the six conditions $(\alpha a)^4(\alpha b)^4(\alpha y) \equiv 0$ are each counted twice. Thus there are $2002 - 7.256 + 21.6 = 364$ independent spreads $(\alpha y)^9 = 0$, and therefore there are $560 - 364 = 196$ cubic spreads in S_{13} which contain $M_5(x)$. We have proved above that all of these are obtained by multiplying each of the 14 quadrics on $M_5(x)$ by the 14 independent linear forms in turn.

As pointed out in **2** the modular manifold $M_5(x)$ contains a significant set of 35 "median points," which map binary octavics with a four-fold root [cf. **2** (9)]. If these equal roots are t_1, t_2, t_7, t_8 or t_3, t_4, t_5, t_6, we find from (5) that $\sigma_1 = \sigma_2 = \sigma_3$. In order that the number may not be greater than 35, the other 12 coördinates σ must also be equal; whence

(23) *The* 35 *median points on* $M_5(x)$ *are associated with the* 35 *triads of* σ's, *and have coördinates*

$$\sigma_1, \sigma_2, \sigma_3, \sigma_4, \cdots, \sigma_{15} = -4, -4, -4, 1, \cdots, 1.$$

A set of five of these median points whose triads exhaust the fifteen σ's [such a distribution as occurs in (11)] will be linearly related and will lie in an S_3. The 56 S_3's of this kind lie on $M_5(x)$, and map binary octavics with a triple root [cf. **2** (11)]. The equations of such an S_3 have the form:

$$(24) \qquad \sigma_1 = \sigma_2 = \sigma_3; \quad \sigma_4 = \sigma_5 = \sigma_6; \quad \cdots; \quad \sigma_{13} = \sigma_{14} = \sigma_{15};$$

the triads in each case being linear. It is easy to verify that these S_3's are contained in the cubic spread (15).

According to **2** (13) the octavics with a double root are mapped upon one of the 28 M_4's, each the section of $M_5(x)$ by an S_8 in the S_{13}. If this double root is $t_7 = t_8$, there follows from **8** (1), (3) that $a = b = \cdots = f = 0$. These six dependent linear relations are expressed in terms of the σ's by the first of equations (20).

With the algebraic character of the invariants (A) thus completely determined for the cases $p = 2$ and $p = 3$, we turn to the similar determination for generic p. This is embodied in the theorem:

(25) *The modular manifold $M_{2p-1}(x)$ in $S_{\nu-1}$* [cf. **2**] *is defined by the aggregate of relations, each quadratic in the linear invariants* (A), *of the form*

$$\frac{(13)(24)(i_5 i_6)\cdots(i_{2p+1} i_{2p+2})}{(14)(23)(i_5 i_6)\cdots(i_{2p+1} i_{2p+2})} = \frac{(13)(24)(j_5 j_6)\cdots(j_{2p+1} j_{2p+2})}{(14)(23)(j_5 j_6)\cdots(j_{2p+1} j_{2p+2})}.$$

For, according to (7) and (22), the linear invariants, when subject to the quadratic conditions, are sufficient to define ordered octavics. By the same argument as was used above in connection with (7), the quadratic relations ensure that these octavics are reunited into an ordered binary $(2p + 2)$-ic.

10. Linear and cubic relations among the invariants (B). The birational relation between the modular manifolds $M_5(x)$ and $M_5(\xi)$. The determination of the algebraic character of the linear invariants (A) given in the preceding section is relatively simple. The invariants (B), though dually related to the invariants (A) [cf. **5**], are of degree 12 in the differences of the roots. Thus it is hardly to be expected that the relations of higher degree satisfied by them would be as simple as the quadratic relations satisfied by the invariants (A) of degree 4 in the differences. They may be regarded as the linear invariants of a set of eight points in S_3, P_8^3, in the particular case when P_8^3 is a set of points on a cubic norm-curve, N^3 [cf. **1** (10)]. But we shall show in a later paper that the 14 linearly independent invariants B must satisfy a system of quintic relations in order that they may define a set P_8^3 (with 9 rather than 13 absolute constants); that they must satisfy a further system of quartic relations in order that they may define a set P_8^3 which is the self-associated set of eight base points of a net of quadrics (with 6 rather than 9 absolute constants); and finally that an additional system of quartic relations must be satisfied in order that the self-associated P_8^3 may be on an N^3 (with 5 rather than 6 absolute constants). We shall find however in the present section a system of cubic relations satisfied by the invariants (B) whose place in the above system of quartic relations is to be discussed later.

We shall use the notation of [**3** (2), **5** (18)] for the 35 determinant products (B),

$$d_{ijk} = \xi_{ijk} = \epsilon_{ijklmno}(ijk8)\,(lmno), \tag{1}$$

where $\epsilon_{ij\ldots o}$ is the sign of the permutation $ij \cdots o$ from the natural order $12 \cdots 7$. These products are connected by 56 linear relations of the two types [cf. **3** (3), (4)]:

$$\begin{aligned} r_{67} &\equiv d_{167} + d_{267} + d_{367} + d_{467} + d_{567} = 0; \\ r_{567} &\equiv d_{567} + d_{234} + d_{134} + d_{124} + d_{123} = 0. \end{aligned} \tag{2}$$

By means of these the number of products which are linearly independent is reduced to 14.

We examine the sum,

$$\tau_1 = d_{127} + d_{347} + d_{567} + d_{135} + d_{146} + d_{236} + d_{245}, \tag{3}$$

and find that it is invariant under the same $g_{8.168}$ as $\sigma_1 = [ad—]$ in **8** (6). It is therefore one of a set, $\tau_1, \cdots, \tau_{15}$, conjugate under $g_{8!/2}$ whose members are permuted cogrediently with $\sigma_1, \cdots, \sigma_{15}$, i. e., like the points of a finite space $S_3(2)$ under the collineation group of the space. There is also a complementary set, $\bar\tau_1, \cdots, \bar\tau_{15}$, which arise from $\tau_1, \cdots, \tau_{15}$ by the transposition (78), and a change of sign. Since there are but 35 d_{ijk}, each must occur in three of the τ's, and also in three of the $\bar\tau$'s, and thus the linear triads in $S_3(2)$ are again encountered.

One such triad is:

$$\begin{aligned} \tau_1 &= d_{127} + d_{347} + d_{567} + d_{135} + d_{146} + d_{236} + d_{245}, \\ \tau_2 &= d_{127} + d_{547} + d_{367} + d_{235} + d_{246} + d_{156} + d_{143}, \\ \tau_3 &= d_{127} + d_{647} + d_{537} + d_{265} + d_{243} + d_{136} + d_{145}. \end{aligned} \tag{4}$$

On adding these, and applying the relations,

$$r_{34} = r_{35} = r_{36} = r_{45} = r_{46} = r_{56} = 0,$$

the sum becomes

$$3d_{127} - 3(d_{456} + d_{356} + d_{346} + d_{345}).$$

This, by virtue of the relation $r_{127} = 0$ is $6d_{127}$; whence

$$6d_{127} = 6\xi_{127} = 6(1278)\,(3456) = \tau_1 + \tau_2 + \tau_3. \tag{5}$$

Again, we prove directly that

$$x_{127} = \left\{ \begin{matrix} 127 \\ 3456 \end{matrix} \right\}^4 = 2(\sigma_1 + \sigma_2 + \sigma_3) \quad [\text{cf. } \mathbf{5}\ (18)]. \tag{6}$$

For, the $\sigma_1, \sigma_2, \sigma_3$ contain 21 invariants (A). Of these the 9 which occur in the formulae **9** (5) have a zero sum both in $\sigma_1 + \sigma_2 + \sigma_3$ and in $\bar{\sigma}_1 + \bar{\sigma}_2 + \bar{\sigma}_3$. The remaining 12 terms in $\sigma_1 + \sigma_2 + \sigma_3$ are the terms in the polar $\left\{ \begin{matrix} 127 \\ 3456 \end{matrix} \right\}^4$ which are odd with respect to the order 3456; the remaining 12 terms in $\bar{\sigma}_1 + \bar{\sigma}_2 + \bar{\sigma}_3$ are the even terms of the polar $-\left\{ \begin{matrix} 127 \\ 3456 \end{matrix} \right\}$. According to **9** (4) the two sets of 12 terms are equal, and each is one half of the polar. The equations (5), (6) give the expressions for the variables x_{ijk}, ξ_{ijk} of the preceding paper in terms of the variables σ, τ we are now using. It is to be observed that the relations (2) among the 35 determinant products ξ_{ijk}, and also among the 35 polars x_{ijk}, are consequences of a single relation $\Sigma_{15}\tau_i = 0$, $\Sigma_{15}\sigma_i = 0$ respectively; e. g.,

$$\begin{aligned}(\tau_1 + \tau_2 + \tau_3) &+ (\tau_4 + \tau_5 + \tau_6) \\ &+ (\tau_7 + \tau_8 + \tau_9) + (\tau_{10} + \tau_{11} + \tau_{12}) + (\tau_{13} + \tau_{14} + \tau_{15}) \\ &\qquad = (d_{127} + d_{137} + d_{147} + d_{157} + d_{167}) = 6r_{17} = 0.\end{aligned}$$

The invariants (B) are cubic polynomials in the invariants (A). For, as a result of **9** (5),

$$5^3 \cdot d_{127} = -(\sigma_1 + \bar{\sigma}_1)(\sigma_1 + \bar{\sigma}_2)(\sigma_1 + \bar{\sigma}_3).$$

This reduces, due to **9** (4), (6), to

$$(7) \qquad 5^3 \cdot d_{127} = -\sigma_1\sigma_2\sigma_3 - \bar{\sigma}_1\bar{\sigma}_2\bar{\sigma}_3.$$

On replacing the $\bar{\sigma}$'s from **8** (11), and reducing the result by using **9** (9), (10), we find that

$$(8) \quad \begin{aligned}10^3 \cdot d_{127} = &-(\sigma_1^3 + \sigma_2^3 + \sigma_3^3) + [\sigma_1^2(\sigma_2 + \sigma_3) \\ &+ \sigma_2^2(\sigma_3 + \sigma_1) + \sigma_3^2(\sigma_1 + \sigma_2)] - 2\sigma_1\sigma_2\sigma_3 + r_1r_2r_3.\end{aligned}$$

That the invariants (B) for generic p can be expressed as polynomials of degree p in the invariants (A) has been proved by Huber.[11]

There follows from (8) that τ_1 can be expressed as a cubic polynomial in the σ's. In this expression the value σ_1 must be isolated since it belongs to the same group as τ_1. Hence the invariants of $g_{8!/2}$ must be subdivided further with reference to the $g_{8.168}$ of σ_1. The list of invariants of $g_{8.168}$ up to the degree three is:

$\sigma_1 = \sigma_1$; $\Sigma_{14}\sigma_2 = -\sigma_1$;
$\sigma_1^2 = \sigma_1^2$; $\Sigma_{14}\sigma_2^2 = -\sigma_1^2 - 2s_2$; $\sigma_1\Sigma_{14}\sigma_2 = -\sigma_1^2$;
$A_1 = \Sigma_7\sigma_2\sigma_3 = 2\sigma_1^2 + s_2/3$; $A_2 = \Sigma_{3.28}\sigma_2\sigma_4 = -\sigma_1^2 + 2s_2/3$;
$\sigma_1^3 = \sigma_1^3$; $\Sigma_{14}\sigma_2^3 = -\sigma_1^3 + 3s_3$; $\Sigma_{14}\sigma_1^2\sigma_2 = -\sigma_1^3$;

$\Sigma_{14}\sigma_1\sigma_2{}^2 = -\sigma_1{}^3 - 2\sigma_1 s_2$; $B_1 = \Sigma_7\sigma_2\sigma_3(\sigma_2+\sigma_3)$; $B_2 = \Sigma_{3.28}\sigma_2\sigma_4(\sigma_2+\sigma_4)$;
$L' = \Sigma_7\sigma_1\sigma_2\sigma_3 = 2\sigma_1{}^3 + \sigma_1 s_2/3$; $L'' = \Sigma_{28}\sigma_2\sigma_4\sigma_6 = -2\sigma_1{}^3 - \sigma_1 s_2/3 + 2s_3$;
$T' = \Sigma_{28.8}\sigma_2\sigma_4\sigma_5$; $T'' = \Sigma_{28}\sigma_3\sigma_5\sigma_7$; $T''' = \Sigma_{28.3}\sigma_2\sigma_3\sigma_5$;
$T^{\mathrm{IV}} = \Sigma_{28.3}\sigma_1\sigma_2\sigma_4 = -\sigma_1{}^3 + 2\sigma_1 s_2/3$.

In this subdivision A_1, A_2 represent, in the finite $S_3(2)$, pairs of points whose join is, or is not, on σ_1. The same distinction appears in B_1, B_2, and in L', L''. The value of A_1 is obtained from the quadratic relation **9** (12); that of A_2 then follows from $s_2 = \Sigma_{14}\sigma_1\sigma$ $A_1 + A_2$. The value of L' is obtained from $\sigma_1 A_1$; that of L'' from $L' + L'' = s_{3L} = 2s_3$ [cf. **9** (18)]. The $T', \cdots, T^{\mathrm{IV}}$ represent sums of products corresponding to triangles in $S_3(2)$ which respectively are on a plane not containing σ_1, are on a plane on σ_1 but with no side on σ_1, have a side on σ_1 but no vertex at σ_1, and have a vertex at σ_1. The value of $T^{(\mathrm{IV})}$ is obtained from $\sigma_1 A_2$. The values of B_1, B_2, T', T'', T''' are still to be determined.

The symmetric function, $\Sigma_{15.14}\sigma_1{}^2\sigma_2 = -3s_3$, yields

$$B_1 + B_2 - 2\sigma_1{}^3 - 2\sigma_1 s_2 = -3s_3.$$

The value $s_{3T} = T' + T'' + T''' + T^{(\mathrm{IV})} = -s_3$ [cf. **9** (18)] yields

$$T' + T'' + T''' - \sigma_1{}^3 + 2\sigma_1 s_2/3 = -s_3.$$

The expression $A_1\Sigma_{14}\sigma_2$ yields

$$B_1 + T''' = -2\sigma_1{}^3 - \sigma_1 s_2/3.$$

The expression $A_2\Sigma_{14}\sigma_2$ yields

$$B_2 + 3L'' + 3T' + 3T'' + 2T''' = \sigma_1{}^3 - 2\sigma_1 s_2/3,$$

an equation which is dependent upon the three which precede it. If the quadratic relation **9** (9) be multiplied by $(\sigma_2 + \sigma_3)$, and the result summed for the seven lines on σ_1, a new equation is obtained:

$$-B_1 + 3T' - 8\sigma_1{}^3 + 8\sigma_1 s_2/3 + 3s_3 = 0.$$

Other summations yield results which are dependent on these. We cannot then obtain all of the invariants of the third degree *integrally* in terms of σ_1, s_2, s_3, though, according to the Galois theory, they can be expressed rationally in terms of $\Delta^{1/2}$, σ_1, and the $s_2, \cdots, s_{15}$. We isolate therefore the 28 terms T'' corresponding to triangles whose planes are on σ_1, but whose sides are not, and find that

$$(10)\qquad \begin{aligned} 2B_1 &= 3T'' - \sigma_1{}^3 - 5\sigma_1 s_2/3,\\ 2B_2 &= -3T'' + 5\sigma_1{}^3 + 17\sigma_1 s_2/3 - 6s_3,\\ 2T' &= T'' + 5\sigma_1{}^3 - 7\sigma_1 s_2/3 - 2s_3,\\ 2T''' &= -3T'' - 3\sigma_1{}^3 + \sigma_1 s_2. \end{aligned}$$

The formulae (8), formed for the seven lines on σ_1, when summed, yield

(11) $$10^3 \cdot \tau_1 = -18\sigma_1^3 - 6\sigma_1 s_2 + 4s_3 + 6T''.$$

This is a pseudo-Tschirnhaus transformation from σ_i to τ_i. It differs from a proper transformation of such sort in that the part T'' in the constant term will vary with σ_i.

If, in the S_{13} of the modular manifold $M_5(x)$, the σ's, with $\Sigma_{15}\sigma_i = 0$, are taken as point coördinates to replace the x_{ijk} of **5** (18), and the τ's, with $\Sigma_{15}\tau_i = 0$, as dual coördinates to replace the ξ_{ijk} of **5** (18), then the incidence condition, $\Sigma_{ijk}x_{ijk}\xi_{ijk} = 0$, becomes, according to (5) and (6),

$$\Sigma_{35}(\sigma_1 + \sigma_2 + \sigma_3)(\tau_1 + \tau_2 + \tau_3) = 0,$$

the summation being extended over the 35 linear triads in $S_3(2)$. This reduces to

(12) $$\Sigma_{15}\sigma_i\tau_i = 0.$$

When, in this condition, the τ's are expressed in terms of the σ's by (11), there results

$$-6\Sigma_{15}\sigma_1^4 - 2s_2\Sigma_{15}\sigma_1^2 + 2\Sigma_{15}\sigma_1 T'' = -6\Sigma_{15}\sigma_1^4 + [\Sigma_{15}\sigma_1^2]^2 + 2\Sigma_{15}\sigma_1 T'' = 0.$$

This takes the simpler form,

(13) $$Q = -5\Sigma_{15}\sigma_1^4 + 2\Sigma_{15.7}\sigma_1^2\sigma_2^2 + 8\Sigma_{15.7}\sigma_1\sigma_2\sigma_4\sigma_9 = 0,$$

where the first two summations are symmetric in the σ's, and the last runs over the 15.7 sets of four points in the finite $S_3(2)$ which lie in a plane and form a base.

Point σ and space τ determined by the same ordered binary octavic are point, and tangent space at the point, of the quartic spread Q. This follows from the fact that τ_1 is the polar of the reference point

$$\sigma_1 = -14, \qquad \sigma_2 = \cdots = \sigma_{15} = 1.$$

For,

$$\partial Q/4\partial\sigma_1 = -5\sigma_1^3 + \sigma_1\Sigma_{14}\sigma_2^2 + 2T'',$$

and $\Sigma_{15}\partial Q/4\partial\sigma_1 = -18s_3 + 2s_{3T} = -20s_3$ [on M_5 (cf. **9** (18))].

The polar of the reference point is thus

$$\tfrac{1}{4}[-14(\partial Q/\partial\sigma_1) + \Sigma_{14}(\partial Q/\partial\sigma_2)] = \tfrac{1}{4}[-15(\partial Q/\partial\sigma_1) + \Sigma_{15}(\partial Q/\partial\sigma_1)]$$
$$= 90\sigma_1^3 + 30\sigma_1 s_2 - 20s_3 - 30T'' = -5.10^3 \cdot \tau_1 \quad \text{[cf. (11)]}.$$

That Q contains $M_5(x)$ is clear from (12) and the theorem of **5** (18). Since the determinant products, and therefore also the τ_i, all vanish simply for a

triple root of the octavic, and doubly for a four-fold root, and since the τ_i are polar cubics of Q, Q must contain the points corresponding to such octavics to a multiplicity two and three respectively. Hence

(14) *The quartic spread Q in* (13) *contains $M_5(x)$. It has triple points at the* 35 *median points* [*the spaces $\{S\}_0$ of* **2** (14)] *which map binary octavics with a four-fold root; and double points throughout the* 56 *S_3's* [*the spaces $\{S\}_1$ of* **2** (14)] *which map binary octavics with a triple root. The point x and space ξ which map the* SAME *octavic are point σ of Q, and tangent space τ of Q at σ.*

In the case $p = 2$ the modular manifold $M_3(x)$ is an $M_3{}^4$ in S_4 and the Tschirnhaus transformation from x to ξ is that of space ξ tangent to $M_3{}^4$ at point x. In the present case $p = 3$ the rôle of $M_3{}^4$ is shared by the cubic spread **9** (15), on which $M_5(x)$ is the nodal locus, and by the quartic spread Q above. The Tschirnhaus property of Q would also obtain for any member of the linear system defined by Q and the cubic spread **9** (15).

We derive finally a set of 15 cubic relations satisfied by the τ's, beginning with the following evident identity connecting the differences of the roots of the octavic:

$$(15) \qquad (1234)(5678)\cdot(1256)(3478)\cdot(1278)(3456) = \Delta^{1/2}[(12)(34)(56)(78)]^2.$$

Expressed in terms of the σ's, τ's by the use of (5), **9** (5), and **8** (11), this is

$$(16) \qquad (\tau_1+\tau_2+\tau_3)(\tau_1+\tau_{10}+\tau_{14})(\tau_1+\tau_5+\tau_7)/6^3 = \Delta^{1/2}(\sigma_1+\bar{\sigma}_1)^2/5^2 = \Delta^{1/2}(\sigma_1-\sigma_2-\sigma_3-\sigma_5-\sigma_7-\sigma_{10}-\sigma_{14})^2/10^2.$$

There are 105 such relations which correspond in the finite $S_3(2)$ to a point σ_1, τ_1 and incident plane $\bar{\sigma}_1, \bar{\tau}_1$ containing three lines on σ_1, τ_1. The 105 squares on the right, quadrics in the σ's, are linearly independent as polynomials in the σ's, but as modular functions they are subject to the 15 linear relations which arise from the quadratic relations R_{ad} of **9** (12). Thus the 105 cubic polynomials in the τ's on the left of (16) must be subject to 15 linear relations. In order to obtain them explicitly, it is necessary to express the quadratic relation **9** (12) in terms of the squares on the right of (16).

Let

$$A = \Sigma_7(\sigma_1-\sigma_2-\sigma_3-\sigma_5-\sigma_7-\sigma_{10}-\sigma_{14})^2,$$

the sum being extended over the seven $\bar{\sigma}$'s on σ_1 in $S_3(2)$; let

$$B = \Sigma_{42}(\sigma_2-\sigma_1-\sigma_3-\sigma_5-\sigma_7-\sigma_{10}-\sigma_{14})^2,$$

the sum being extended over the 42 points not at σ_1 on the seven planes containing σ_1 in $S_3(2)$; and let

$$C = \Sigma_{56}(\sigma_2 - \sigma_4 - \sigma_5 - \sigma_6 - \sigma_8 - \sigma_{10} - \sigma_{13})^2,$$

the sum being extended over the 56 points in the 8 planes not on σ_1 in $S_3(2)$. Then, making use of (9) in which the value of A_1 is obtained from the quadratic identity **9** (12), we find that

$$(17)\quad A = 20\sigma_1^2 - 8s_2/3,\quad B = 20\sigma_1^2 - 88s_2/3,\quad C = -40\sigma_1^2 - 48s_2,$$
$$4A - 2B + C = 0.$$

When similar summations, $A(\tau)$, $B(\tau)$, $C(\tau)$, are made on the left members of (16) [the numerical factor $1/6^3$ being disregarded], and the results expressed in terms of sums similar to those occurring in (9), we find that

$$A(\tau) = 4\tau_1^3 + L''(\tau) + T''(\tau) + T^{(IV)}(\tau),$$
$$B(\tau) = 3\tau_1\Sigma_{14}\tau_2^2 + 3\Sigma_{14}\tau_2^3 + 3B_1(\tau) + B_2(\tau) + 12L'(\tau) + 3L''(\tau)$$
$$+ 3T''(\tau) + 4T'''(\tau) + 3T^{(IV)}(\tau),$$
$$C(\tau) = 4\Sigma_{14}\tau_2^3 + 2B_2(\tau) + 8L''(\tau) + 4T'(\tau).$$

The combination, $4A(\tau) - 2B(\tau) + C(\tau)$, then yields the cubic relation satisfied by the τ's:

$$16\tau_1^3 - 6\tau_1\Sigma_{14}\tau_2^2 - 2\Sigma_{14}\tau_2^3 - 6B_1(\tau) - 24L'(\tau) + 6L''(\tau)$$
$$+ 4T'(\tau) - 2T''(\tau) - 8T'''(\tau) - 2T^{(IV)}(\tau) = 0.$$

The following relations, most of which are consequences of $\Sigma_{15}\tau_i = 0$, may be used to modify the cubic relation:

$$L'(\tau) + L''(\tau) + T'(\tau) + T''(\tau) + T'''(\tau) + T^{(IV)}(\tau) = s_3(\tau),$$
$$B_1(\tau) + B_2(\tau) = 2\tau_1^3 + 2\tau_1 s_2(\tau) - 3s_3(\tau),$$
$$(18)\quad L'(\tau) + T^{(IV)}(\tau) = \tau_1^3 + \tau_1 s_2(\tau),$$
$$\tau_1\Sigma_{14}\tau_2^2 = -\tau_1^3 - 2\tau_1 s_2(\tau),$$
$$\Sigma_{14}\tau_2^3 = -\tau_1^3 + 3s_3(\tau).$$

A convenient form of the cubic relation is

$$(19)\quad R_1(\tau) \equiv 12\tau_1^3 + 6\tau_1 s_2(\tau) - 4s_3(\tau) - 3B_1(\tau)$$
$$- 11L'(\tau) + 4L''(\tau) + 3T'(\tau) - 3T''(\tau) \equiv 0.$$

It is easily verified that $\Sigma_{15}R_1(\tau) \equiv 0$, as should be the case, since for the σ's, $\Sigma_{15}R_i = 0$.

These cubic relations take a more elegant form when the three relations (19) attached to the three points τ of a line in $S_3(2)$ are added. The square

on the right of (15) corresponds in the $S_3(2)$ to an "element," i. e., point σ_1 and *incident* plane $\bar{\sigma}_1$. When a line in $S_3(2)$ is given with 3 incident points p and 3 incident planes π, the 105 elements divide into 9 of type πp, 12 of type $\pi p'$, 12 of type $\pi' p$, and 72 of type $\pi' p'$. When the identities (17) determined by the three points p are added, a square determined by an element of type πp enters into the sum with a coefficient $4 - 2 - 2 = 0$, one of type $\pi p'$ with a coefficient $-2 - 2 - 2 = -6$, one of type $\pi' p$ with a coefficient $4 + 1 + 1 = 6$, and one of type $\pi' p'$ with a coefficient $-2 + 1 + 1 = 0$. Hence the derived identity expresses that the sum of one set of 12 squares equals the sum of the other set. For the line in $S_3(2)$ isolated by the division of indices 1357, 2468, the identity among the squares takes the form,

$$(20) \qquad \Sigma\, [(1i)(3j)(5k)(7l)]^2 - \Sigma\, [(1j)(3i)(5k)(7l)]^2 = 0,$$

the summation being extended over the 12 even permutations $ijkl$ of 2468. By adding the seven relations (20) attached to the seven lines on a point in $S_3(2)$, the relation (17) is again obtained.

If the squares in (20) are expressed in terms of the determinant products as in (15), we have, on noting the change of sign in $\Delta^{1/2}$ under an odd permutation, the following cubic identity among the invariants (B):

$$(21) \qquad R_{246} \equiv \Sigma d_{ij2} d_{ik4} d_{il6} = 0 \qquad (i, j, k, l = 1, 3, 5, 7).$$

There are 35 relations like (21) but they all are linear combinations of 14 of the relations (19). The question as to whether these relations completely characterize the invariants (B) will be discussed later.

URBANA, ILLINOIS.

REFERENCES

[1] A. B. Coble, "Point sets and allied Cremona groups," *Transactions of the American Mathematical Society*, I: Vol. 16 (1915), pp. 155-198.

[3] A. B. Coble, "Algebraic geometry and theta functions," *Colloquium Publications of the American Mathematical Society*, New York, Vol. 10 (1929).

[8] A. B. Coble, "Hyperelliptic functions and irrational binary invariants," *American Journal of Mathematics*, Vol. 54 (1932), pp. 425-452.

[9] M. Noether, "Ueber die Gleichung achten Grades und ihr Auftreten in der Theorie der Curven vierter Ordnung," *Mathematische Annalen*, Vol. 15 (1879), pp. 89-110.

[10] E. H. Moore, "Concerning the general equations of the seventh and eighth degrees," *Mathematische Annalen*, Vol. 51 (1899), pp. 417-444.

[11] C. M. Huber, "On complete systems of irrational invariants of associated point sets," *American Journal of Mathematics*, Vol. 49 (1927), pp. 251-267.

GROUPS WHOSE ORDERS INVOLVE A SMALL NUMBER OF UNITY CONGRUENCES.

By G. A. Miller.

For the sake of brevity we shall use the term *unity congruence* of the natural number g to represent the property that a prime factor of g is congruent to unity with respect to another such factor, and the sum of the numbers of such congruences for the various prime factors of g will be called the number of its unity congruences. It will at first be assumed that g is not divisible by the square of a prime number and the number of the possible groups of order g will be determined when g involves a small given number k of unity congruences. It is well known that a necessary and sufficient condition that there is one and only one group of order g is that g involves no unity congruence.

A very useful theorem in the determination of the possible groups of order g when g involves a given number of unity congruences is that if two or more distinct sets of factors of g can be found which are such that no prime factor of one of these sets is congruent to unity with respect to a prime factor of another then every group of order g is the direct product of subgroups whose orders are the products of the prime factors in these sets.* Another theorem which is frequently useful in this determination may be stated as follows: *If g involves exactly k unity congruences and k prime factors of g are congruent to unity with respect to the same prime factor p of g then the number of the groups of order g is*

$$1 + \Sigma c_i (p-1)^{i-1} \qquad (i = 1, 2. \cdots, k).$$

where c_i represents the number of combinations of k things taken i at a time.

A proof of this theorem results directly from the facts that every such group is the direct product of a group whose order is the product of the given $k+1$ prime factors and a cyclic group whose order is the product of the remaining prime factors of g and that the quotient group of the former with respect to its largest Sylow subgroup is cyclic. To an operator of prime order in this quotient group there must correspond an operator of the same order in the group and a necessary and sufficient condition that in an automorphism of the group such an operator can correspond to a

* G. A. Miller, *Proceedings of the National Academy*, Vol. 18 (1932), p. 472.

power of itself whose index is not congruent to unity with respect to its order is that it is commutative with the operators of the given largest Sylow subgroup. If p is congruent to unity with respect to the prime number r, but no other prime factor of g has this property nor is r congruent to unity with respect to a prime factor of g then the number of groups of order gr is just one more than the number of the groups of order g.

From the given theorems it results directly that when g involves k unity congruences and also at least $2k$ prime factors then it is always possible to select g so that there are exactly 2^k groups of order g which are separately the direct product of a cyclic group and of k groups each of whose orders involves exactly 2 prime factors. When $k > 1$ all of these groups must be of odd order but when $k = 1$ the groups may be of even order, and this is then the only case which can present itself since it is assumed that g is not divisible by the square of a prime number. The cyclic group in question reduces to the identity when and only when g involves only $2k$ prime factors, and there is always one and only one such direct product which is abelian. This is also cyclic since every abelian group whose order is not divisible by the square of a prime number is cyclic.

When $k = 2$ and a prime factor of g is congruent to unity with respect to a second such factor which is itself thus congruent with respect to a third such factor then it results from a general theorem * that there are exactly 3 groups of order g. If a prime factor of g is congruent to unity with respect to two other such factors then it results from another known theorem * that there are exactly 4 groups of order g. Finally, when two distinct prime factors of g are congruent to unity with respect to the same third prime factor p it results from the theorem noted above that there are exactly $p + 2$ groups of order g. As the case when g involves at least four prime factors such that two are separately congruent to unity with respect to the other two respectively was considered in the preceding paragraph, we have established the following theorem: *If we know only that the order g of a group is not divisible by the square of a prime number but that it involves exactly two unity congruences then one of four and only four cases is possible. In one of these there are exactly 3 groups of order g, in each of two others there are exactly 4 such groups, while in the remaining case the number of the possible groups is $p + 2$, where p is a prime number. All of these groups are of odd order except that in the last case they are of even order if and only if $p = 2$.*

We proceed to consider the possible groups when $k = 3$ and g is not

* G. A. Miller, *Proceedings of the National Academy*, Vol. 18 (1932), p. 472.

divisible by the square of a prime number. From the theorem just noted it results that when such a group is the direct product of a group whose order involves exactly one unity congruence and a group whose order involves exactly two such congruences then the number of the possible groups of order g is one of the following three numbers: 6, 8, $2p+4$. As direct products are well known when their constituent factors are known and as the number of the unity congruences of the order of such a group, which is the direct product of groups such that the order of no one of the factor groups involves a prime number which is congruent to unity with respect to a prime factor of the order of another factor group, is the sum of the unity congruences of the orders of these factor groups we may confine our attention in what follows to groups which are not such direct products.

If g involves a prime factor which is congruent to unity with respect to three other such factors then g must be odd and there are 8 groups of this order. If g involves a prime factor which is congruent to unity with respect to two other such factors and one of these two is thus congruent with respect to the other p there are exactly $p+4$ such groups as was noted by O. Hölder for the special case when g involves only three prime factors.* If one and only one of these two factors is thus congruent with respect to a fourth prime factor of g there are exactly 6 groups of order g and g must be odd. If this fourth factor is thus congruent with respect to the first of the three given prime factors there are exactly 5 groups of order g since the two subgroups whose orders are this fourth prime factor and the first such factor respectively generate an invariant subgroup of such a group.

It remains to examine the cases when no prime factor of g is congruent to unity with respect to as many as two other such factors. If a prime factor of p is congruent to unity with respect to a second such factor and this is thus congruent with respect to a third which is again thus congruent with respect to a fourth then it results from a theorem to which we referred above that there are exactly 5 groups of order g. When the first of the given prime factors is congruent to unity with respect to the second and this second is thus congruent with respect to a third the fourth must be thus congruent either to the second or to the third and the number of the groups of order g is then $p+3$ in the former case according to the theorem noted above in the second paragraph, while it is $p+4$ in the latter case, where p is the prime factor with respect to which this fourth prime factor of g is thus congruent.

It remains to examine the cases when if one prime factor of g is con-

* O. Hölder, *Mathematische Annalen*, Vol. 43 (1893), p. 412.

gruent to unity with respect to another such factor this cannot be thus congruent with respect to a third such factor. When three prime factors of g are congruent to unity with respect to a fourth such factor p then it results from the theorem noted in the second paragraph above that the number of the groups of order g is $p^2 + p + 2$, where g may be so chosen that p is any prime number including 2. When two prime factors of g are congruent to unity with respect to a third such factor then none of these factors could be congruent to unity with respect to a fourth such factor in accord with the cases excluded above nor could this fourth be congruent to unity with respect to any one of these three. Hence all the possible cases have been considered and the following theorem is established: *A natural number g can be so chosen that it involves exactly three unity congruences but is not divisible by the square of a prime number and that the number of groups of order g is exactly equal to an arbitrary one of the following numbers* 5, 6, 8, $p + 3$, $p + 4$, $2p + 4$, $p^2 + p + 2$, *where p is a prime number, but g cannot be so selected that the number of these groups is not included in this list.*

In the considerations which follow it is desirable to use certain new extensions of Sylow's theorem which will now be explained. It is well known that G. Frobenius extended in 1895 Sylow's theorem by proving that the number of the subgroups of order p^a, p being a prime number, in any group whose orders is divisible by this number, is congruent to unity with respect to p. When p^a is the highest power of p which divides this order then all of these subgroups are conjugate under the group but for lower powers of p there may be more than one complete set of such conjugate subgroups. In a permutation group of degree n the number of these sets which are such that the subgroups of this order contained therein include only subgroups which are conjugate under the symmetric group of degree n can, however, not exceed the number of such sets composed separately of subgroups which are conjugate under the normaliser of a Sylow subgroup of order p^m of the group. Hence the following theorem: *If the order of a permutation group of degree n is divisible by p^a, p being a prime number, then the number of the different sets of conjugate subgroups of order p^a involving only subgroups of this order which are conjugate under the symmetric group of degree n can not exceed the number of such sets composed separately of subgroups contained in a given Sylow subgroup and conjugate under the normaliser thereof. It can also not be less than the number of the latter sets which are composed separately of conjugates under this symmetric group.*

As a special case of this theorem there results the abstract group theorem that the number of the different sets of conjugate subgroups of order p^a which

involve only simply isomorphic subgroups of this order and appear in a given group can not exceed the number of the sets composed of simply isomorphic groups which appear in a given Sylow subgroup of order p^m and are separately composed of the subgroups of order p^a which are conjugate under the normaliser of this Sylow subgroup. As a still more special theorem it may be noted that all the subgroups of order p^a contained in a group whose Sylow subgroups of order p^m are cyclic are contained in a single set of conjugate subgroups under the group.

As an illustrative example of this theorem it may be noted that the octic group is a Sylow subgroup of the group of order 168 and of degree 7. This Sylow subgroup involves three and only three subgroups of order 4, which are invariant under it. One of these subgroups is cyclic, another is transitive, and the third is intransitive. Hence the simple group of order 168 contains exactly three sets of conjugate subgroups of order 4. From the cited abstract group theorem it results that the number of these sets of conjugate subgroups is either two or three. The relation between the possible number of conjugate subgroups of a group and of one of its Sylow subgroups is perhaps more fully illustrated by the subgroups of order 2 contained in the octic group.

It is well known that this group contains three sets of conjugate subgroups of order 2 and hence it follows from the given theorem that if this group is a Sylow subgroup of a given group then the number of its sets of conjugate subgroups of order 2 cannot exceed three. As a matter of fact this number is exactly three in the group of order 72 and of degree 6. It is two in the octahedral group, while it is only one in the simple group of order 168. Hence all the possibilities allowed by the given theorems actually appear in these various groups.

The fact that the simple group of order 168 contains exactly two complete sets of conjugate non-cyclic subgroups of order 4 results also from another abstract group theorem which we proceed to develop. Since every subgroup of index p is invariant under a group of order p^m it results that the number of the subgroups of order p^{m-1} which appear in a complete set of conjugates of a group involving a Sylow subgroup of order p^m is always prime to p. When $p = 2$ this number is therefore odd. A dihedral group of order 2^m contains exactly three subgroups of index 2 and when $m > 2$ two of these subgroups are non-cyclic and simply isomorphic. Hence it results that if a group involves a Sylow subgroup of order 2^m, $m > 2$, which is dihedral then the number of its cyclic subgroups of order 2^a, $\alpha > 1$, must always be odd and hence the number of its non-cyclic subgroups of order

2^{m-1} must always be even. That is, these simply isomorphic non-cyclic subgroups can not appear in a single complete set of conjugates. This proves the following theorem: *If a group contains a dihedral Sylow subgroup of order 2^m, $m > 2$, then its cyclic subgroups of order 2^α, $\alpha > 1$, appear in a single complete set of conjugates but its non-cyclic subgroups of order 2^{m-1} appear in two such sets.*

When $m > 3$ it is obvious that a similar theorem applies to the groups which involve the dicyclic group of order 2^m as a Sylow subgroup. It also results directly from these considerations that if a group involves the quaternion group as a Sylow subgroup then its subgroups of order 4 appear either in a single set of conjugates or in three such sets. That is, there can not be exactly two complete sets of conjugate subgroups of order 4 in a group which has the quaternion group as a Sylow subgroup. If a group involves an abelian Sylow subgroup of order 2^m then the number of its complete sets of conjugate subgroups of the same order which is a power of 2 must always be odd since the order of the normaliser of such a subgroup must be divisible by 2^m. The sum of the indexes of the normalisers of a set of subgroups of order p^a which is composed of one and only one from each complete set of conjugate subgroups of this order is always congruent to unity with respect to p. In particular, if a group has a Sylow subgroup of order 8 then it involves exactly three complete sets of conjugate subgroups of order 4 when this Sylow subgroup is the octic group or the abelian group of type $(2, 1)$. It involves one, three, five, or seven such sets when this subgroup is the abelian group of type $(1, 1, 1)$. When this subgroup is cyclic it involves just one such set and when it is quaternion it involves either one or three such sets as was noted above.

From the theorems just noted it results that when g is divisible by the square of one and only one prime number p and involves no unity congruences then the number of groups of order g is $1 + 2^k$, where k is the number of the prime factors of g which divide $p + 1$. When g involves one and only one unity congruence and is equal to $p_1^2 p_2 \cdots p_\lambda$, where $p_1, p_2, \cdots, p_\lambda$ are distinct prime numbers then every group of order g involves an invariant subgroup of order p_1^2 or of order p_2. If p_1 is congruent to unity with respect to another prime factor of g this fact results directly from the proof of a well-known theorem relating to groups whose orders are not divisible by the square of a prime number. If p_2 is thus congruent then it results directly from the same theorem that every group of order g contains exactly $p_1^2 p_2$ operators whose orders divide this number. Hence the theorem in question results directly from the transformations of a subgroup of order p_2 under the opeartors of a group of order p_1^2.

When one of the factors of g is 2 then g must be either of the form $2p^2$ or of the form $4p$, where p is an odd prime number. In the former case there are 5 groups of order g and this is also the number of such groups in the latter case when either $p = 3$ or $p - 1$ is divisible by 4, while there are exactly 4 such groups when neither of these conditions is satisfied. It remains to consider the cases when g is odd. If $p_1 - 1$ is divisible by a prime factor p of g but $p_1 + 1$ does not have this property there are exactly $p + 4$ groups of order g since an operator of order p which transforms into itself the non-cyclic group of order p^2 must transform into themselves at least two of the $p_1 + 1$ subgroups of order p in this non-cyclic group and if it transforms into themselves more than two such subgroups it must transform each of these subgroups into itself. If $p_1 - 1$ is divisible by a prime factor p and $p_1 + 1$ is divisible by k distinct prime factors of g the number of the groups of order g is therefore $2 + (2 + p)2^k$.

If p_2 is congruent to unity with respect to p_1 but not with respect to p_1^2 and if k of the prime factors of g divide $p_1 + 1$ there are exactly $3 + 2$ groups of order g, while there is one more such group when p_2 is thus co gruent to unity with respect to p_1^2. When p_2 is thus congruent with respec to another prime factor of g the number of these groups is $2 + 2^{k+1}$. Hence the following theorem among others has been established. *If the natural number g is divisible by the square of one and only one prime number p but by no higher power of this prime and if g involves no unity congruence the number of the possible groups of order g is $1 + 2^k$, where k is the number of the prime factors of g which divide $p + 1$.*

COMPLEMENTS OF POTENTIAL THEORY. PART II.*

By Griffith C. Evans.

1. *Introduction.* In general form, as a statement of Gauss's theorem in the plane, Poisson's equation may be expressed by the relation

$$\int_s D_n u \, ds = \Phi(s), \tag{1.1}$$

* Presented to the American Mathematical Society, September, 1931. Literature will be cited as follows:

(I) G. C. Evans, "Sopra un'equazione integro-differenziale di tipo Bôcher," *Rendiconti della R. Accademia dei Lincei*, Vol. 28 (1919), pp. 262-265.

(I') J. Radon, "Über die Randwertaufgaben beim logarithmischen Potential," *Sitzungsberichte der Akademie der Wissenschaften in Wien*, Vol. 128 (1919), pp. 1123-1167.

(II) G. C. Evans, "Fundamental points of potential theory," *Rice Institute Pamphlet*, Vol. 7 (1920), pp. 252-329.

(III) G. Vitali, "Analisi delle funzioni a variazione limitata," *Rendiconti del Circolo Matematico di Palermo*, Vol. 46 (1922), pp. 388-408.

(IV) F. Riesz, "Über subharmonische Funktionen und ihre Rolle in der Funktionentheorie und in der Potentialtheorie," *Acta Univ. Franc. Jos. Szeged*, Vol. 2 (1925), pp. 87-100.

(V) A. J. Maria, "Functions of plurisegments," *Transactions of the American Mathematical Society*, Vol. 28 (1926), pp. 448-471.

(VI) G. C. Evans, *The Logarithmic Potential. Discontinuous Dirichlet and Neumann Problems*, New York (1927).

(VII) J. E. Littlewood, "Mathematical notes (7); on functions subharmonic in a circle," *Journal of the London Mathematical Society*, Vol. 2 (1927), pp. 192-196.

(VIII) J. E. Littlewood, "Mathematical notes (8); on functions subharmonic in a circle," *Proceedings of the London Mathematical Society*, Vol. 28 (1928), pp. 383-394; reported November, 1927.

(IX) G. C. Evans, "Discontinuous boundary value problems of the first kind for Poisson's equation," *American Journal of Mathematics*, Vol. 51 (1929), pp. 1-18; preliminary report presented to the American Mathematical Society, September, 1927.

(X) E. R. C. Miles, "Boundary value problems for potentials of a single layer," *Transactions of the American Mathematical Society*, Vol. 31 (1929), pp. 190-203.

(XI) G. C. Evans and E. R. C. Miles, "Potentials of general masses in single and double layers. The relative boundary value problems," *Proceedings of the National Academy of Sciences*, Vol. 15 (1929), pp. 102-108; *American Journal of Mathematics*, Vol. 53 (1931), pp. 493-516.

(XII) F. Riesz, "Sur les fonctions subharmoniques et leur rapport à la théorie du potential," *Acta Mathematica*, Vol. 54 (1930), pp. 321-360. This paper is a sequel to a paper of the same title, *ibid.*, Vol. 48 (1926), pp. 329-343.

(XIII) G. C. Evans, "Complements of potential theory," *American Journal of Mathematics*, Vol. 54 (1932), pp. 213-234.

where $D_n u$ is the generalized or vector derivative of u in the direction of the interior normal,* and $\Phi(s)$ is a completely additive function of curves or plurisegments s.† For many purposes it is convenient to have $\Phi(s)$ a function with regular discontinuities; we assume therefore that $\Phi(s)$ is given by the equation

$$(1.2) \qquad \Phi(s) = \int q(s, P)\, d\Phi(e_P),$$

where $\Phi(e)$ is the mass on the set e (measurable Borel) for an arbitrarily given distribution of finite positive and negative mass on a bounded open set T, of the plane, where $q(s, P)$ is the symmetric (circular) density at P of the region σ enclosed by s, and the integration is extended over the entire plane. Poisson's equation (1.1) is assumed to hold for almost all curves of a class of simple smooth rectifiable curves, that is to say, except for those which contain on their arcs portions, of positive linear measure, of some exceptional set of superficial measure zero.

As for the class of curves s, one may limit oneself to rectangles, or to the whole or a subclass of curves of class Γ; ‡ a curve of this class is a simple

---

The author should have been earlier familiar with (I′), in which J. Radon gives a thoroughgoing analysis, by means of the Fredholm theory as extended by F. Riesz and himself, of the *continuous* Dirichlet problem, and of the *generalized* Neumann problem in the case where the potential of the single layer on the boundary C is itself continuous on C. The memoir should have been cited in (XI), which deals with the *generalized discontinuous* Dirichlet and Neumann problems, by reduction of the Stieltjes integral equations to the usual Fredholm type. J. Radon considers less general distributions on the boundary, even in the case of the Neumann problem. than the later papers of the present author; on the other hand, his advance in the treatment of the equations in linear operators enables him to extend radically the type of boundary to which the essential ideas of the method of integral equations apply. In this sense it completes the study by Plemelj, *Potentialtheoretische Untersuchungen*, Leipzig (1911), (cited in (VI), p. 55), as the papers of the author do in the direction of general mass distributions.

With Radon, the boundary is one of "bounded turning." The angle turned through by the tangent to C as the tangent progresses around C is a function of bounded variation of the arc. Accordingly a denumerable infinity of vertices is allowed, subject to the condition of bounded total turning. The possible finite number of cusps is ruled out for the sake of obtaining suitable inequalities on the special parameter values in the integral equations.

* See Appendix I.

† The theory of such functions is given in (V), which is a generalization of (III).

‡ For applicability to the logarithmic potential, consideration of the general class Γ demands an extension of the definition of $\int_s u\, dx$ (See (II), p. 264). No such extension is necessary if the class is restricted to a subclass of Γ for which $\int_s \log 1/QP\, ds_P$ exists.

closed curve composed of a finite number of arcs with continuously turning tangent, for which there is a constant Γ such that

$$\int_s \frac{|\cos(n_P, MP)|}{MP} \, ds_P < \Gamma, \quad \int_s \frac{|\cos(n_Q, QP)|}{QP} \, ds_P < \Gamma,$$

for all M in the plane and all Q on s.

In an earlier treatment of the Dirichlet problem for Poisson's equation, by means of the Green's function and conformal transformations,* the author asserted that for sufficiently smooth boundaries such problems could be handled by direct methods, which, at the same time, would be applicable to three dimensions. It is the purpose of the present paper to make this extension, and to treat also the discontinuous Neumann problem, in which is given the limit of the flux $\int D_n u \, ds$ over an arc which approaches an arc of the boundary.

Let C be a simple closed curve in the plane, which may or may not pass through points of T or contain them in its interior region. Let Σ, Σ' be the interior and exterior regions, respectively, defined by C. The curve is to have a normal at every point, the symbol n denoting the direction towards the interior. The normals at P, Q on C are to satisfy the uniform Neumann condition

$$(\alpha) \qquad |\sphericalangle (n_P, n_Q)| < \alpha s,$$

where s is the shorter of the arcs PQ and α is a positive constant. The curve C is therefore a curve of class Γ.†

Consider also a class $\{C'\}$ of simple closed rectifiable curves C', neighboring C, with normal at every point, such that if n_P is the normal to C at P on C, and n'_R the normal to C' at R on C', we have

$$(\beta) \qquad |\sphericalangle (n_P, n'_R)| < \alpha \overline{PR}.$$

There is no loss of generality in assuming the same constant α in both inequalities. If C' is close enough to C there is defined a one-to-one correspondence of points of C and points of C' along normals to C. For a given C' we let τ be the upper bound of these normal distances, and consider an ordered set of these C' such that τ approaches zero.

A particular family $\{C'\}$ is that constituted by curves parallel to C, that is, by curves C' whose normal distance from C along normals to C is constant and equal to τ, τ small enough.

That the potential

* (IX). † (XI).

$$(1.3) \qquad V(M) = \int_T \log \frac{1}{r}\, d\Phi(e_P), \qquad r = MP,$$

is a solution of Poisson's equation for almost all s was proved by the author in 1920,* also that it is a potential function for its generalized derivative †

$$(1.4) \qquad D_\alpha V = \int_T \frac{\cos(\alpha, r)}{r}\, d\Phi(e_P).$$

We wish to obtain further properties of $V(M)$, and prove first the following theorem.‡

THEOREM I. *As M approaches Q on C along n_Q,*

$$\lim_{M=Q} V(M) = V(Q)$$

for almost all Q.

In fact, $V(Q)$ evidently exists almost everywhere on C, and is summable on C. In order to prove the theorem, however, we have need of a lemma.

2. *Lemma on integrals over $\{C'\}$.* Consider a set of non-overlapping intervals (A_i, B_i) on C, in number m, and corresponding intervals (A'_i, B'_i) on various curves C'_i of $\{C'\}$, cut off by normals to C, and form the integral

$$I = \Sigma_i \int_{A_i}^{B_i} \{V(M) - V(Q)\}\, ds_Q$$

where Q is on (A_i, B_i), M the corresponding point on (A'_i, B'_i) and ds_Q is the element of arc of C.

LEMMA. *Given $\epsilon > 0$ we can find $\tau > 0$ so that if all the τ_i are $\leq \tau$ we shall have $|I| < \epsilon$, irrespective of m and of the curves C'_i on which the intervals corresponding to the (A_i, B_i) are given.*

It is sufficient to prove the lemma for $\Phi(e)$ of positive type, that is, $\Phi(e) \geq 0$ for all e (meas. B). We have

$$(2.1) \qquad I = \Sigma_i \int_{A_i}^{B_i} ds_Q \int_T \log(QP/MP)\, d\Phi(e_P).$$

* (II).

† See (II), also Appendix I.

‡ The corresponding fact for the case of the unit circle, that $\lim\,(r = 1)\, U(r, \theta) = 0$ for almost all θ, where $U(r, \theta) = U(M) = \int g(M, P)\, d\phi(e_p)$, $g(M, P)$ being the Green's function for the circle with pole at M, was proved independently and simultaneously by J. E. Littlewood (VIII) and the author (IX), and as F. Riesz has pointed out results also from the analysis of subharmonic functions (XII, 1930). The theorem given in the text has wider generality.

That $\int_s |\log QP|\, ds_Q$, $\int_s |\log MP|\, ds_Q$ and also $\int_{s'} |\log MP|\, ds'_M$ converge even when QP or MP passes through zero during the integration may be easily verified, taking τ sufficiently small. The fact will be established in the course of the proof. But granting it, for the moment, we may utilize a well known theorem and interchange the order of integration in (2.1), and write

$$(2.2) \qquad I \leq \int_T d\Phi(e_P)\Sigma_i \int_{A_i}^{B_i} |\log(QP/MP)|\, ds_Q.$$

For convenience we divide T into two portions T', T'', T'' being the portion of T which lies between two curves parallel to C and distant from it by a small amount τ', and T' being the remainder of T. Let I', I'' be the corresponding parts of I. We suppose τ to be $< \tau'$.

We also divide the intervals $\Sigma(A_i, B_i)$, $\Sigma(A'_i, B'_i)$ into two portions, for the sake of treating I'', denoting by p_1, p_2, p'_1, p'_2 the parts of these sets, as follows. Let P_0 be the foot of the normal to C through P, which is in T'', and let C_δ be the portion of C of which P_0 is the center, of length 2δ. Let p_1 be the portion of $\Sigma(A_i, B_i)$ which lies in C_δ, p_2 the rest of it. Similarly let p'_1 be the portion of $\Sigma(A'_i, B'_i)$ whose projection by normals to C lies in C_δ, p'_2 the rest of it.

We may suppose τ' and δ to be small enough so that the following relations are valid, for P in T'', M in p'_1, Q in p_1:

$$\begin{aligned} &PM \geq \text{arc } P_0Q/2, \\ &PQ \geq \text{arc } P_0Q/2, \qquad QP < 1,\ MP < 1, \\ &2ds_Q \geq ds'_M \geq ds_Q/2. \end{aligned}$$

These relations follow in an obvious manner from the conditions (α), (β) and the fact that two normals to C at points in C_δ, δ small enough, cannot intersect in such a way that the length of *either* to the point of intersection will be $\leq 1/\alpha$.* Thus we may write

---

* At any point P_0 of C draw two circles, each of radius $1/\alpha$, tangent to C at P_0; then C_δ lies between these two circles. In fact, if x, y are rectangular coördinates, P_0 being the origin and the common tangent being the x axis, if P is on C and P' on the one of the two circumferences whose center is on the positive y axis, such that arc P_0P' = arc P_0P, we have, for $x_P > 0$,

$$x_P = \int_{P_0}^{P} \cos(x, s)\, ds > \int_{P_0}^{P'} \cos(\alpha s)\, ds = x_{P'},$$

$$|y_P| = \left| \int_{P_0}^{P} \sin(x, s)\, ds \right| < \int_{P_0}^{P'} \sin(\alpha s)\, ds = y_{P'}.$$

This geometric situation holds therefore for tangent circles at any point of C, the circles having the given radius.

3

$$\int_{p_1} |\log QP|\, ds_Q + \int_{p_1} |\log MP|\, ds_Q \leq 4 \int_{-\delta}^{\delta} |\log P_0 Q|\, ds_Q \leq m(\delta)$$

where $m(\delta)$ is a quantity independent of P in T'' which approaches zero with δ.

For P in T'' and M in p'_2, we have Q in p_2, $PM \geq \delta/2$, $PQ \geq \delta/2$. Hence, with τ', δ fixed, by making τ small enough, less than say τ'', we can make $|\log(QP/MP)|$ uniformly less than ϵ_2, given $\epsilon_2 > 0$. Consequently, denoting meas. C by C, we shall have

$$|I''| < [m(\delta) + C\epsilon_2]\, \Phi(T'').$$

Also, for the consideration of I', by making τ small enough, say $< \tau'''$, we can make $|\log(QP/MP)| < \epsilon_3$, given arbitrarily $\epsilon_3 > 0$, while the previously given quantities remain fixed. For P is in T'; hence τ may be chosen small enough so that MP, QP remain bounded away from zero as τ tends to zero. Finally, then,

$$|I| < [m(\delta) + C\epsilon_2]\, \Phi(T) + C\epsilon_3 \Phi(T).$$

Having fixed τ', we can therefore take δ and then τ small enough so that $|I| < \epsilon$. But this proves the lemma.

Incidentally we notice that *the same inequality applies to* (2.1) *when* $\log(QP/MP)$ *is replaced by its absolute value, and also to the integral* I_1*:*

$$(2.3) \qquad I_1 = \Sigma_i \int_{A'_i}^{B'_i} V(M)\, ds'_M - \Sigma_i \int_{A_i}^{B_i} V(Q)\, ds_Q.$$

For $\lim(\tau = 0)$ [arc PM/arc P_0Q] $= 1$, uniformly, again by condition (β). The curve C may or may not be a depository for mass itself, and the curves C' may or may not cross C. Nor need the curve C be closed, although we have taken it that way for the sake of the applications.

3. *Proof of Theorem I and quasi continuity of* $V(M)$. It is sufficient to give the proof for $\Phi(e)$ of positive type. In this case $V(M)$ is lower semi-continuous in the finite plane.* Accordingly, with the lemma, the con-

---

Construct a symmetric neighborhood on C, about P_0, of suitably small length 2δ. Suppose now, contrary to what we wish to prove, that at P on C in this neighborhood, say for $x > 0$, the normal cuts n_{P_0} at a point M such that $P_0M \leq 1/a$. This yields a contradiction, for the construction of the two circles, of radius $1/a$, tangent to C at P shows that C cannot go through P_0. In fact, let K be the point on PM, or PM produced, which is the center of one of these two circles. If K is interior to the segment PM we have $P_0K < P_0M \leq 1/a$. If K is beyond or at M, $P_0K \leq P_0M + MK < PM + MK = 1/a$. Hence in both cases $P_0K < 1/a$, and the circle with center M and radius $1/a$ contains P_0 in its interior.

* (IV), p. 98; (IX), p. 6.

ditions of the theorem and corollary of Appendix II are satisfied, and $\lim(M = Q \text{ on } n_Q)\ V(M) = V(Q)$, for almost all Q on C. But this is Theorem I.

The following corollary to Theorem I comes incidentally, merely as a special case of the lemma.

COROLLARY. *Let τ approach zero, M'_1, M'_2 on C' approach Q_1, Q_2 on C; then* $\lim \int_{M=M'_1}^{M=M'_2} V(M)\, ds_Q = \int_{Q_1}^{Q_2} V(Q)\, ds_Q$ *and* $\lim \int_{M'_1}^{M'_2} V(M)\, ds'_M = \int_{Q_1}^{Q_2} V(Q)\, ds_Q$, *for all Q_1, Q_2 on C.*

In fact, on account of the uniformity of the inequalities on I and I_1 and the continuity of the integral over arcs of C, it is not necessary for M'_1 M'_2 to lie on n_{Q_1}, n_{Q_2}, respectively.

As an application of Theorem I, let C be a rectangle of which one side $x = x_0$ is AB. By Theorem I, $\lim(x = x_0)\ V(x,y) = V(x_0, y)$, for almost all y in AB. Similarly $\lim(y = y_0)\ V(x,y) = V(x, y_0)$ for almost all x. Nevertheless this function need not be continuous as a point function in the plane at any point whatever in T; in fact $V(M)$ may be infinite at all the rational points of T while it remains the potential of a finite mass. We shall prove however the following theorem.

THEOREM II. *For almost all y, $V(x,y)$ is an absolutely continuous function of x, and for almost all x, an absolutely continuous function of y.*

For the proof, we may limit ourselves to $\Phi(e)$ of positive type. Since $v(M) = v(x,y) = \int_T (1/MP)\, d\Phi(e_P)$ is summable superficially it is summable as a function of x for almost all y; in fact the double integral of this positive function is convergent, and by a fundamental theorem may be expressed as an iterated integral. Let t denote a line $y = \text{const.}$ corresponding to any one of these non-exceptional values of y.

We note that *the total mass on t vanishes, that is, if e is any bounded set. meas. B, we have* $\Phi(e \cdot t) = 0$. In fact,

$$\int_T \frac{1}{MP}\, d\Phi(e_P) = \int_{T-t} \frac{1}{MP}\, d\Phi(e'_P) + \int_t \frac{1}{MP}\, d\Phi(e''_P),$$

where $e'_P = e_P \cdot (T - t)$, $e''_P = e_P \cdot t$, both of the integrals of the right hand member being convergent almost everywhere on t and representing summable functions of x. We write

$$\begin{aligned} h_n(M,P) &= 1/MP, & MP &\geqq 1/n \\ &= n, & MP &< 1/n. \end{aligned}$$

Then, by definition of the generalized integral and well known properties of the Lebesgue integral, we have

$$\int_{x_0}^{x_1} dx_M \int_t \frac{1}{MP}\, d\Phi(e''_P) = \int_{x_0}^{x_1} dx_M \,[\lim_{n=\infty} \int_t h_n(M,P)\, d\Phi(e''_P)]$$

$$= \lim_{n=\infty} \int_{x_0}^{x_1} dx_M \int_t h_n(M,P)\, d\Phi(e''_P)$$

$$= \lim_{n=\infty} \int_t d\Phi(e''_P) \int_{x_0}^{x_1} h_n(M,P)\, dx_M.$$

But this last quantity may be evaluated directly, because

$$\int_{x_0}^{x_1} h_n(M,P)\, dx_M \geq 1 + \log n - \log \frac{x_1 - x_0}{3}, \quad \text{if} \quad x_0 \leq x_P \leq x_1,$$

for sufficiently large n. If we denote this right hand member by N, we have $\lim(n=\infty)\, N = \infty$. Hence our iterated integral, assumed to be convergent by hypothesis, satisfies the inequality

$$\int_{x_0}^{x_1} dx_M \int_t \frac{1}{MP}\, d\Phi(e''_P) \geq N\Phi(t'), \quad N \textit{ arbitrarily large},$$

where t' is the closed interval (x_0, x_1) on t. Consequently $\Phi(t') = 0$, $\Phi(t) = 0$ and $\Phi(e \cdot t) = 0$.

With this fact established, the convergence of $\int_{x_0}^{x_1} dx \int_T (1/MP)\, d\Phi(e_P)$ enables us to interchange the orders of integration and evaluate the inside integral which we obtain by substituting $\cos(x, MP)/MP$ for $1/MP$. We have, writing $r = MP$, $M_1 = (x_1, y)$, $r_1 = M_1P$, etc.,

$$(3.1) \quad \int_{x_0}^{x_1} dx \int_T \frac{\cos(x,r)}{r}\, d\Phi(e) = \int_{x_0}^{x_1} dx \int_{T-t} \frac{\cos(x,r)}{r}\, d\Phi(e)$$

$$= \int_{T-t} d\Phi(e) \int_{x_0}^{x_1} \frac{\cos(x,r)}{r}\, dx = \int_{T-t} \{\log(1/r_1) - \log(1/r_0)\}\, d\Phi(e)$$

$$= \int_T \{\log(1/r\,) - \log(1/r_0)\}\, d\Phi(e),$$

for all x_0, x_1.

But for any y, the integral defining $V(M)$ is convergent for almost all x, in particular, for x_0 properly chosen. It follows, by adding the integral expression for $V(x_0, y)$ to the last member of (3.1) that the resulting integral is convergent; but this is $V(x, y)$. Hence, for all x_0, x_1,

$$(3.2) \qquad \int_{x_0}^{x_1} dx \int_T \frac{\cos(x,r)}{r}\, d\Phi(e) = V(x_1, y) - V(x_0, y).$$

For these non-exceptional values of y, the function $V(x, y)$ is therefore absolutely continuous as a function of x, and the partial derivative of $V(x, y)$ exists for almost all x, and has the value

$$\frac{\partial V}{\partial x} = \int_T \frac{\cos(x, r)}{r} \, d\Phi(e).$$

But, since $V(x, y)$ is measurable superficially and continuous in x for almost all y, the subset of T where the derivative fails to exist is measurable superficially, and must accordingly have measure zero.* Moreover, almost everywhere in T the generalized derivative $D_x V$ exists and has the value just given for the partial derivative. Hence the two are equal almost everywhere. Similar reasoning applies to $V(x, y)$ considered as a function of y. Thus we have established Theorem II, and also, incidentally, the following corollary.

COROLLARY. *The partial derivatives of $V(x, y)$ exist and are identical with the generalized derivatives almost everywhere.*

4. *$V(M)$ a function of class (ii).* We have established the fact that $\lim(M = Q) V(M) = V(Q)$ for almost all Q on C, approach being along normals, and also that if $\{C'_i\}$ is a denumerable sequence of curves C'_i approaching C in such a way that τ_i approaches zero, the integral $\int_{(C'_i)} V(M) ds_Q$, extended over C'_i, approaches $\int_C V(Q) ds_Q$. The same remarks apply if for V is substituted $V + K$, where K is an arbitrary constant. Consider then $\Phi(e)$ of positive type, and K large enough so that $V(M) + K$ is positive in a closed region which contains in its interior the set T, the curves C'_i and the curve C; in fact, $V(M)$ is bounded below, in such a region, being $\geq \log(1/d)\Phi(T)$, where d is the diameter of the region or unity, according as to which is the greater.

It follows from De la Vallée Poussin's converse of Vitali's theorem †

---

* That the partial derivatives of $V(x, y)$ exist almost everywhere is stated by F. Riesz (IV), who mentions that it follows from the Lebesgue theory, and by Evans (IX) without proof. However, if in (3.1) we except by hypothesis the denumerable infinity of lines $y = \text{const.}$, on each of which there may be a positive mass, we have $\int_T = \int_{T-t}$, and the statement is an immediate consequence of (II), § 1.4, p. 259. F. Riesz mentions also the existence almost everywhere of the generalized Laplacian in Petrini's form (H. Petrini, *Acta Mathematica*, Vol. 31 (1908), at p. 181. See also Evans, *Cambridge Colloquium Lectures*, New York (1918), at p. 85), with reference to Green's theorem.

† C. de la Vallée Poussin, "Sur l'intégrale de Lebesgue," *Transactions of the American Mathematical Society*, Vol. 16 (1915), pp. 435-501; see p. 445.

that the absolute continuity of $\int \{V(M) + K\}\, ds_Q$ and therefore of $\int V(M)\, ds_Q$ (and also, of course, of $\int |V(M)|\, ds_Q$) is uniform over $\{C'_i\}$. The same remark accordingly applies if $\Phi(e)$ is the difference of two set functions of positive type, that is, if $\Phi(e)$ is an arbitrary additive function of point sets. Moreover for ds_Q may be substituted ds'_M on account of the relations given for those quantities in § 2. Thus we establish the following theorem:

THEOREM III. *If $\{C'_i\}$ is a denumerable sequence of curves satisfying (β) which approaches C, satisfying (α), then the absolute continuity of the integral $\int V(M)\, ds_Q$ is uniform over $\{C'_i\}$.*

COROLLARY. *The absolute continuity of $\int V(M)\, ds'_M$ is uniform over $\{C'_i\}$, ds'_M being the element of arc of C'_i.*

This uniform property may be extended to other families of curves C', not necessarily denumerable. We shall speak of a *normal* family $\{C\}$ of curves C if each curve of the family satisfies (α), if every two curves of the family satisfy (β) with respect to each other, the constant being fixed for the family, and if in every infinite subset of the family there is a subsequence $\{C'_i\}$ which approaches a curve C of the family, approach being in the sense that the maximum normal distance of C_i from C approaches zero. Such a family is, for example, that of the curves parallel to C in its neighborhood.

THEOREM IV. *The absolute continuity of $\int V(M)\, ds_M$ is uniform over all curves C of a normal family.*

In fact, otherwise there is a sequence of curves of the family over which the absolute continuity is not uniform, and therefore a denumerable subsequence of the same kind which approaches some C of the family so that τ approaches zero. But this contradicts the previous theorem.

The author has used the term *class* (ii)* to denote functions $u(M)$ which satisfy the uniform absolute continuity property with respect to families of concentric circles all inside or all outside a given one, or, for curves other than circles, a corresponding property in terms of conformal transformation. The term may be extended to the present situation without ambiguity.

A function $u(M)$ will be said to be of *class* (ii) inside (or outside) a curve C which has the property (α) if the uniform absolute continuity

* (VI), (IX).

property of Theorem IV holds for every normal family which includes C, but otherwise lies entirely inside (or outside) C.

A function $u(M)$ will be said to be of *class* (i), inside (or outside) C, if $\int_C |u(M)|\, ds_M$ remains bounded for similar normal families. The class (ii) is a subclass of (i).

THEOREM V. *The function $V(M)$ is of class (ii) and of class (i).*

5. *The discontinuous Dirichlet problem.* The difference of two functions which are of class (ii), or of class (i), inside (or outside) C, and which are potential functions of their generalized derivatives and solutions of Poisson's equation (1.1) will be a function of the same kind, which is a solution inside (or outside) C of the equation

$$\int_s D_n u \, ds = 0. \tag{5.1}$$

But such a function, by a generalization of Bôcher's theorem, has merely unnecessary discontinuities; and when these are removed, by changing the value of the function at most on a set of superficial measure zero, the function becomes continuous with all its derivatives, inside (or outside) C, and satisfies Laplace's equation.* Hence the solution of the interior (or exterior) discontinuous boundary value problems for (1.1), where $u(Q)$ is given as a summable function on C, or where the average value of $u(Q)$ is given on C (that is, $\lim(\tau = 0) \int_{Q_1}^{Q_2} u(M)\, ds_M = F(Q_2) - F(Q_1)$, where $F(Q_2) - F(Q_1)$ is the value for the segment Q_1Q_2 of a given additive function of plurisegments with regular discontinuities) is reduced to the solution of the corresponding problems for Laplace's equation.

These discontinuous Dirichlet problems for Laplace's equation have been solved for surfaces subject to the condition (α), in terms of potentials of double layers on the boundary,† with the customary modification for the exterior problem, and that analysis is valid in two dimensions. It will be shown in a subsequent note that these solutions of Laplace's equation belong to the classes (ii) or (i) respectively, according as potentials are chosen so that the limiting values of $u(M)$ or the limiting values of the average of $u(M)$ may be given on C, and that they are the only solutions in these respective classes which fit the boundary values, for the interior and exterior problems.

---

* (II), § 6.3. See also Evans, "Note on a theorem of Bôcher," *American Journal of Mathematics*, Vol. 50 (1928), pp. 123-126.

† (XI).

6. *The generalized Neumann problem.* A similar method may be used with regard to the generalized Neumann problem, where the limiting values of the flux are given for approach to C from inside or from outside C. In order to handle this problem, we shall need the condition (j).* We say that a function $u(M)$ belongs to the *class* (j), inside (or outside) C_0, if the quantity

$$\int_C |D_n u(M)|\, ds_M$$

is bounded *on almost all curves* of any normal family of curves C which contains C_0 but otherwise lies entirely inside (or outside) C_0.

We shall prove the following theorem:

THEOREM VI. *Let $\{C\}$ constitute a normal family in the neighborhood of C_0, lying entirely inside (or outside) C_0, whose curves approach C_0, as τ approaches zero, from inside* [$\lim \tau = 0+$] (*or from outside* [$\lim \tau = 0+$]) *C_0. Let Q_1, Q_2 be given on C_0, M_1, M_2 the corresponding points on C. Then except for those curves of $\{C\}$ which contain on their arcs portions of positive linear measure of a certain exceptional set of superficial measure zero, the quantities $\int_{M_1}^{M_2} D_n V\, ds_M$, $\int_{M_1}^{M_2} |D_n V|\, ds_M$ are defined and bounded as τ approaches zero, and*

$$\lim_{\tau=0\pm} \int_{M_1}^{M_2} D_n V\, ds_M = \mp \pi\{\mu(Q_2) - \mu(Q_1)\} + \int_{Q_1}^{Q_2} ds_Q \int_T \frac{\cos(QP, n_Q)}{QP}\, d\Phi(e_P), \tag{6.1}$$

where $\mu(Q)$ is the function of bounded variation, with regular discontinuities, which measures the mass distribution on C_0 itself.

Incidentally, $V(M)$ is of class (j).

We may limit ourselves to $\Phi(e)$ of positive type. Except on a certain set of superficial measure zero we have

$$D_\alpha V = \int_T \frac{\cos(MP, \alpha)}{MP}\, d\Phi(e_P)$$

for any direction α. Hence on almost all C we have

$$\begin{aligned} \int_{M_1}^{M_2} D_n V\, ds_M &= \int_{M_1}^{M_2} ds_M \int_T \frac{\cos(MP, n_M)}{MP}\, d\Phi(e_P) \\ &= \int_T d\Phi(e_P) \int_{M_1}^{M_2} \frac{\cos(MP, n_M)}{MP}\, ds_M \end{aligned} \tag{6.2}$$

* (VI).

the change of order of integration being justifiable.* Similarly

$$\int_{M_1}^{M_2} | D_n V \mid ds_M \leq \int_T d\Phi(e_P) \int_C \frac{|\cos(MP, n_M)|}{MP} ds_M.$$

But on account of the condition (α) the quantity

$$\int_C |\cos(MP, n_M)|/MP \, ds_M$$

is bounded, less than some constant Γ which is independent of C.† Hence on almost all C we have $\int_C | D_n V \mid ds_M \leq \Gamma\Phi(T)$, which proves the boundedness of the quantities $\int_{M_1}^{M_2} D_n V \, ds_M$, $\int_{M_1}^{M_2} | D_n V \mid ds_M$ as specified in Theorem VI, and, incidentally, that $V(M)$ is of class (j).

We divide T into three portions, namely, the set of points of T on C_0 itself, which we denote by D, and T' and T''. By T'' we denote the portion of T, except for D, which lies between two curves parallel to C_0 and distant from it by a small amount τ_1; we define $T' = T - T'' - D$. Then

$$\Phi(e) = \Phi(e \cdot D) + \Phi(e \cdot T') + \Phi(e \cdot T''),$$

and the integral

$$I = \int_T d\Phi(e) \int_{M_1}^{M_2} \frac{\cos(MP, n_M)}{MP} ds_M$$

may be written as the sum of three corresponding parts

$$I = I_0 + I' + I''.$$

We have $| I'' | \leq \eta\Gamma$, where $\eta = \Phi(T'')$ may be made arbitrarily small with τ_1, independently of τ; and I' is continuous as τ passes through zero. Consequently

$$\lim_{\tau=0\pm} (I' + I'') = \int_{Q_1}^{Q_2} ds_Q \int_T \frac{\cos(QP, n_Q)}{QP} d\Phi(e_P \cdot [T' + T'']). \tag{6.3}$$

But I_0 is given in terms of a single layer distribution on C_0. If we let $\mu(Q)$ represent the non-decreasing function with regular discontinuities which corresponds to this distribution on C_0, we have ‡

$$\lim_{\tau=0\pm} I_0 = \mp \pi\{\mu(Q_2) - \mu(Q_1)\} + \int_{Q_1}^{Q_2} ds_Q \int_{C_0} \frac{\cos(QP, n_Q)}{QP} d\mu(P). \tag{6.4}$$

If now we add (6.3) and (6.4), writing $d\mu(P) = d\mu(e_P \cdot D)$, we obtain

$$\lim_{\tau=0\pm} I = \mp \pi\{\mu(Q_2) - \mu(Q_1)\} + \int_{Q_1}^{Q_2} ds_Q \int_T \frac{\cos(QP, n_Q)}{QP} d\Phi(e_P),$$

* (II), see § 3. 1. † As in (XI), § 2. ‡ (X), (XI).

which, on account of (6. 2), is (6. 1). We may now return to the general additive function $\Phi(e)$, and this establishes the theorem.

The exceptional curves, where the flux integral is not defined, do not appear if the flux is defined by a limiting process like that employed by F. Riesz.*

The difference of two functions which satisfy Poisson's equation (1. 1), are potential functions of their generalized derivatives inside (or outside) C_0 and are of class (j) inside (or outside) C_0 is a function of the same kind, which except for removable discontinuities satisfies Laplace's equation. It will be shown in a subsequent note that the harmonic functions of class (j) are identical with those which are potentials of a single layer distribution on C_0 itself. But among these potentials there is a unique solution of the boundary value problem †

$$\lim_{\tau=0-}\int_{M_1}^{M_2} D_n V\, ds_M = G(Q_2) - G(Q_1)$$

and of the problem

$$\lim_{\tau=0+}\int_{M_1}^{M_2} D_n V\, ds_M = G(Q_2) - G(Q_1)$$

where $G(Q)$ is a function of bounded variation on C_0, with regular discontinuities, and, for the interior problem, such that

$$\int_C dG(Q) = 0.$$

Appendix I.

ON POTENTIAL FUNCTIONS OF GENERALIZED DERIVATIVES.

7. More than a decade ago,‡ the author employed the notions of *generalized derivative* for functions of several variables, and *potential function of vector or generalized derivatives,* in order to utilize a property analogous

---

* (XII).

† (I′), (X), (XI).

‡ (II), See pp. 274-285. There are one or two rather obvious corrections necessary in the proof of the theorem of § 5. 33, already mentioned in (IX), p. 146, namely:

P. 279, last equation; In order that $\int_{\sigma_1} u(\partial v/\partial x)\,d\sigma$, of the right hand member, should converge as a double integral, it is necessary to state an additional hypothesis on v, say that $\partial v/\partial x$ be bounded.

P. 282, equation (23). Let $\partial v/\partial x$ and $\partial w/\partial y$ similarly be bounded.

P. 283. Let $\partial^2\theta/\partial x^2$, $\partial^2\theta/\partial x\,\partial y$ etc. all be bounded. Otherwise, one may obviate the difficulty by rewriting the integrals like $\int_{\sigma_1} u(\partial v/\partial x)\,d\sigma$ as iterated, instead of double, integrals.

Also, it should be mentioned that equation (25′), p. 285, is true for almost all θ_1, θ_2, as specified in previous equations of this type.

to absolute continuity in the several separate variables, sufficiently general nevertheless to include applications to Newtonian or logarithmic potentials of absolutely general distributions of finite positive and negative mass. For such functions are not necessarily continuous.

Somewhat later, L. Tonelli * formulated the definition of "funzione di due variabili assolutamente continua," in a study of general problems relating to areas of surfaces. In this appendix the two notions are compared. It happens that if in the definition of potential function of generalized derivatives the function is assumed to be continuous, as a point function, the specialized concept thus obtained is identical with the one, just mentioned, formulated by Tonelli.

Let $u(x,y)$, summable superficially over a bounded open region T, be also summable on a class of simple rectifiable curves sufficiently general to include all rectangles in T. Let α be an arbitrary direction and α' the direction $\pi/2$ in advance of α; let s be an arbitrary closed curve of the class, and let σ denote its interior region and the measure of that region. Then the generalized derivative of u in the direction α is defined as the quantity

$$D_\alpha u = \lim_{\sigma=0} \frac{1}{\sigma} \int_s u \, d\alpha' \tag{7.1}$$

where σ shrinks to zero as a regular family, that is, so that σ/d^2 (d being the diameter of σ) remains greater than some positive constant, as d tends to zero.

Suppose now in particular that

$$\Phi_x(\sigma) = \int_s u \, dy, \qquad \Phi_y(\sigma) = -\int_s u \, dx$$

generate absolutely continuous functions of point sets, $\Phi_x(e)$ and $\Phi_y(e)$ respectively, in the interior of any region which, with its boundary lies in T. Then $\int_s u \, d\alpha'$ generates also an absolutely continuous function of point sets $\Phi_\alpha(e)$; in fact

$$\Phi_\alpha(e) = \Phi_x(e)\cos(x,\alpha) + \Phi_y(e)\cos(y,\alpha). \tag{7.2}$$

Therefore, for any α, $D_\alpha u$ exists except on a set of superficial measure zero, independent of α, and has the vector property

$$D_\alpha u = D_x u \cos(x,\alpha) + D_y u \cos(y,\alpha). \tag{7.3}$$

* L. Tonelli, "Sulla quadratura delle superficie," *Rendiconti della R. Accademia dei Lincei* (6), Vol. 3 (1926, 1 sem.), pp. 633-638, and "Sulle funzioni di due variabili assolutamente continue," *Memorie della R. Accademia delle Scienze dell'Istituto di Bologna* (Scienze fisiche), (8), Vol. 6 (1928-29), pp. 81-88.

The quantity $D_\alpha u$ is merely the Lebesgue derivative of the absolutely continuous function of point sets $\Phi_\alpha(e)$. Moreover $D_\alpha u$ is summable over any closed domain interior to T and

$$(7.4) \qquad \int_s u\, d\alpha' = \int_\sigma D_\alpha u\, d\sigma.$$

If, for every direction α, (7.4) *holds for all s, σ* (*of the given class*) *in T, $D_\alpha u$ being, for every fixed α, summable over any region which with its boundary is contained in T, the function u is said to be a potential function of its generalized or vector derivatives in T.*

Obviously, the concept can be extended to any number of independent variables, with surfaces or hypersurfaces instead of curves.

The connection with absolute continuity in the separate variables lies in the fact that in any rectangular portion of T, if (x_0, y_0), (x, y) are taken outside a certain point set of superficial measure zero, the function

$$(7.5) \qquad \bar{u}(x, y) = \int_{y_0}^{y} D_y(x_0, \eta)\, d\eta + \int_{x_0}^{x} D_x(\xi, y)\, d\xi + u(x_0, y_0)$$

is identical with $u(x, y)$.* The same holds for the analogous function in which the rôles of x and y are interchanged.

According to Tonelli, a function $u(x, y)$ *is an absolutely continuous function of two variables in the square* $(0, 1)$ if

(7.6)
- (i) $u(x, y)$ is continuous;
- (ii) $u(x', y)$ is absolutely continuous in y for almost all x', and $u(x, y')$ is absolutely continuous in x for almost all y';
- (iii) the total variation $u_{(y)}(x')$ of $u(x', y)$ is a summable function of x' in $(0, 1)$, and the total variation $u_{(x)}(y')$ of $u(x, y')$ is a summable function of y' in $(0, 1)$.

Similarly the concept is defined for any rectangle. Let us say further, in accordance with the general idea of Tonelli, that the function is absolutely continuous in x and y in T if it is absolutely continuous according to the above definition in every rectangular region which, with its boundary, is contained in T.

8. We show now that (7.6) implies (7.4). From (7.6) the functions $\partial u/\partial x$, $\partial u/\partial y$ are summable over any closed domain in T, since such a domain may be contained in a finite number of closed rectangular regions, contained in T, which have no parts in common but their boundaries. For such a

* (II), p. 278.

rectangle, say (x_1, y_1), (x_2, y_2), contained with its boundary in T, we have, by Fubini's well known theorem

$$\int_\sigma \frac{\partial u}{\partial x}\, d\sigma = \int_{y_1}^{y_2} dy \int_{x_1}^{x_2} \frac{\partial u}{\partial x}\, dx = \int_{y_1}^{y_2} \{u(x_2, y) - u(x_1, y)\} dy = \int_s u\, dy.$$

Similarly for any region σ, composed of a finite number of such rectangles, distinct except for their boundaries,

$$(8.1) \qquad \int_\sigma \frac{\partial u}{\partial x}\, d\sigma = \int_s u\, dy; \qquad \int_\sigma \frac{\partial u}{\partial y}\, d\sigma = -\int_s u\, dx.$$

But $\int_\sigma \frac{\partial u}{\partial x}\, d\sigma$, $\int_\sigma \frac{\partial u}{\partial y}\, d\sigma$ define absolutely continuous functions of point sets in any region which with its boundary is contained in T; moreover $u(x, y)$ is continuous in T. Hence the relations (8.1) hold for any domain in T bounded by a finite number of simple closed rectifiable curves lying in T. Also the Lebesgue derivatives $D_x u$, $D_y u$, of these absolutely continuous functions, exist almost everywhere in T, and almost everywhere have the values $\partial u/\partial x$, $\partial u/\partial y$ respectively.

Consequently (8.1) become

$$(8.2) \qquad \int_\sigma D_x u\, d\sigma = \int_s u\, dy, \qquad \int_\sigma D_y u\, d\sigma = -\int_s u\, dx.$$

But (8.2) imply (7.4), and $u(x, y)$ is a potential function of its generalized derivatives.

9. Conversely, suppose that $u(x, y)$ is a potential function of its generalized derivatives *and is continuous.* Then, from (7.5), in any rectangle in T,

$$u(x, y) = u(x_0, y_0) + \int_{x_0}^{x} D_x u(\xi, y)\, d\xi + \int_{y_0}^{y} D_y u(x_0, \eta)\, d\eta$$

is an identity in x for almost all y; for u and $\bar{u}$ are both continuous in x for almost all y. Hence $u(x, y)$ is absolutely continuous in x for almost all y, and $\partial u/\partial x$ exists and has the value $D_x u$ almost everywhere in T. Finally, $D_x u$ is summable over any rectangle (x_1, y_1), (x_2, y_2), contained with its boundary in T, and therefore $u_{(x)}(y') = \int_{x_1}^{x_2} | \partial u(x, y')/\partial x |\, dx$ is a summable function of y' in (y_1, y_2). Evidently the rôles of x and y may be interchanged, and Tonelli's conditions (7.6) are satisfied. The function $u(x, y)$ is therefore absolutely continuous in x and y.

10. The author has established, for potential functions of generalized derivatives, the identity

$$\int_\sigma \frac{1}{r} D_r u \, d\sigma = \int_{\theta_1}^{\theta_2} \{u(r_2, \theta) - u(r_1, \theta)\} \, d\theta,$$

where σ is the region $r_1 < r < r_2$, $\theta_1 < \theta < \theta_2$, in polar coördinates, and more generally, a similar identity for other curvilinear coördinates.* Thus if u is continuous, it follows as before that

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x} \cos(x, r) + \frac{\partial u}{\partial y} \cos(y, r)$$

almost everywhere in a given circle in T whose center is the origin of polar coördinates, and that u is absolutely continuous as a function of r for almost all θ. Thus it happens that the theorems proved by Tonelli in the last cited reference † are in their essence special cases of those given earlier by the author.

A still less restrictive definition of potential function of generalized derivatives has been given lately by the author.‡ The function $u(x, y)$ is summable, and the quantities $\int_s u \, dy$, $-\int_s u \, dx$ are assumed to be defined merely for "almost all" rectangles, or for "almost all" curves of a certain class which includes all rectangles. If these quantities determine absolutely continuous functions of point sets, in any region contained, with its boundary, in T, $u(x, y)$ is said to be a potential function of its generalized derivatives in T; $D_x u$ and $D_y u$ are the Lebesgue derivatives of these two absolutely continuous functions of point sets. The essential theorems remain valid.

In particular, the theorem of the cited note enables one to pass from Poisson's equation to Laplace's equation: the difference of any two functions which are potential functions of their generalized derivatives (considering merely rectangles), and satisfy Poisson's equation (1.1) for "almost all" rectangles is, except for removable discontinuities, a harmonic function.

The concept which Tonelli has named "quasi-absolutely continuous function of two variables" in order to distinguish it from the "absolutely continuous function of two variables" is intermediate between the latter and the "potential function of generalized derivatives" referred to rectangles. It is defined by the conditions (ii), (iii) of (7.4).§ The function $V(M)$ is a quasi-absolutely continuous function of x and y.

* (II), pp. 282, 287.

† Tonelli, *loc. cit. Bologna.*

‡ "Note on a theorem of Bôcher," *American Journal of Mathematics*, Vol. 50 (1928), pp. 123-126.

§ L. Tonelli, "Sur la quadrature des surfaces," *Comptes Rendus Hebd. des Séances de l'Académie des Sciences*, Vol. 182 (1926), pp. 1198-1200.

APPENDIX II.

A PARTIAL CONVERSE OF VITALI'S THEOREM.

11. Vitali's fundamental theorem may be stated as follows. If $\{f_n(x)\}$ is a sequence of functions, summable $a < x < b$, with $\lim(n = \infty) f_n(x) = f(x)$, and if the absolute continuity of $\int f_n(x)\,dx$ in (a, b) is uniform, then $f(x)$ is summable and $\lim(n = \infty) \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx$. De la Vallée Poussin's converse of this theorem, for a sequence of not negative functions, has been used in § 4, above. The theorem and corollary of this appendix constitute a partial converse of the theorem, but from another point of view.

We consider a function $f(x, y)$, defined in the interval $a < x < b$, for each y, $0 \leq y < c$, which is lower semi-continuous in x for each positive y and summable in x for y equal to zero. We form the expressions

$$I_1 = \sum_1^n{}_i \int_{x'_i}^{x_i''} f(x, y_i)\,dx, \qquad 0 < y_i \leq Y < c,$$

$$I_0 = \sum_1^n{}_i \int_{x'_i}^{x_i''} f(x, 0)\,dx,$$

where the intervals (x'_i, x_i''), contained in the open interval (a, b), are finite in number and non-overlapping (except possibly with end points in common). The theorem is as follows: *

THEOREM. *If, independently of the choice of n and $y_1, y_2, \cdots, y_n$,*

$$\overline{\lim_{Y=0}}\; I_1 \leq I_0,$$

then

$$\overline{\lim_{y=0+}}\; f(x, y) \leq f(x, 0),$$

for almost all x in (a, b).

The hypothesis is that given $\epsilon > 0$, we can choose $Y > 0$ so that for all n, and any set of values $y_1, \cdots, y_n$, $0 < y_i \leq Y$, we have $I_1 < I_0 + \epsilon$.

The set of values of x, corresponding to points (x, y), $0 < y \leq Y$, where $f(x, y)$ is greater than a given number η, is open, since the function is lower semi-continuous in x. But this is the same set as that on which $F(x) > \eta$, where $F(x)$ is the upper bound of $f(x, y)$ considered as a function of y, $0 < y \leq Y$. Accordingly $F(x)$ is measurable, and also $F(x) - f(x, 0)$. Hence the set of values of x, corresponding to points (x, y), $0 < y \leq Y$,

---

* This theorem is an extension of that of (II), p. 317.

where $f(x,y) - f(x,0) > \eta > 0$, is measurable; consequently also the set where

$$\overline{\lim_{y=0}}\, f(x,y) \;\; - f(x,0) > \eta.$$

If then we assume that the theorem is not true, there are positive numbers η, m and a set k_1 of values of x, meas. $k_1 > 2m > 0$, such that

$$\overline{\lim_{y=0}}\, f(x,y) \;\; - f(x,0) \; > 2\eta. \qquad x \text{ in } k_1.$$

In particular, we take $Y > 0$ small enough so that

$$I_1 - I_0 < \tfrac{1}{2}\eta m.$$

Let $f(x)$ be a continuous approximation to $f(x,0)$, which differs from $f(x,0)$ by $< \eta$ except on a set of measure $< m$, and such that

$$\int_a^b |\, f(x,0) - f(x)\,|\, dx < \tfrac{1}{2}\eta m.$$

Such a function is furnished by the average

$$f(x) = f_\mu(x,0) = \frac{1}{2\mu}\int_{x-\mu}^{x+\mu} f(\xi,0)\, d\xi,$$
$$f(\xi,0) = 0 \text{ for } \xi \geqq b \text{ and } \xi \leqq a,$$

taking μ sufficiently small.*

We write

$$I = \sum_1^n{}_i \int_{x'_i}^{x_i''} \{f(x,y_i) - f(x)\}\, dx$$

and it follows that $\qquad I < \eta m$

and $\qquad \overline{\lim_{y=0}}\, f(x,y) - f(x) > \eta,$

on some measurable set k of measure $> m$. We may assume k to be closed since it contains a closed set of measure differing from that of the original set by as little as we choose.

Let E be the set of points (x,y), $0 < y \leqq Y$, for which $f(x,y) - f(x) > \eta$. The projection of E on the x-axis includes k. Moreover, since $f(x)$ is continuous, $f(x,y) - f(x)$ is lower semi-continuous as a function of x, and therefore if this function is greater than η at a point (x,y), there is a neighborhood of (x,y) on the line $y = \text{const.}$ where the function is $> \eta$; hence every point of E is an interior point of some interval (ξ'_i, ξ_i''), $y = \eta_i > 0$, which lies in E.

---

* H. E. Bray, "Proof of a formula for an area," *Bulletin of the American Mathematical Society*, Vol. 29 (1923), pp. 264-270.

It follows therefore that each point of the closed set k is an interior point of one or more of the projections of these intervals. There will then be a finite number of these intervals (ξ'_i, ξ_i'') which cover k. As portions of these we can therefore select a finite number N of closed subintervals (X'_i, X_i'') which do not overlap, except at their extremities, and cover k. The corresponding intervals in E will lie on various lines Y_i, $Y_i \leq Y$. But then

$$I = \sum_1^N {}_i \int_{x'_i}^{x_i''} \{f(x, Y_i) - f(x)\}\, dx > \eta \sum_1^N \int_{x'_i}^{x_i''} dx$$

or $\qquad I > \eta m$

which is a contradiction with the inequality previously obtained. Thus the theorem is proved.

Further, supposing that $f(x, 0)$ is still summable with respect to x, let us assume that, for almost all x, $f(x, y)$ is lower semi-continuous in y at $y = 0$; that is, given almost any x, if $f(x, 0) > N_x$, then $f(x, y) > N_x$ for $y > 0$, small enough. We have therefore, for almost all x, $\lim(y = 0)\ f(x, y) \geq f(x, 0)$. If, however, we make use of the result of the theorem just proved, it follows that $\lim(y = 0)\ f(x, y) = f(x, 0)$.

Corollary. *If, further, for almost all x, $f(x, y)$ is lower semi-continuous in y at $y = 0$, $f(x, y)$ has $f(x, 0)$ as a true limit, for almost all x in (a, b).*

In the application in the text, the situation corresponds to that in which $f(x, y)$ is lower semi-continuous in (x, y) for $y \geq 0$, and $\lim(Y = 0) \mid I_1 - I_0 \mid = 0$, so that the conclusion of this corollary is in force.

We note finally that in the theorem and corollary of this appendix the range of values of y need not be continuous. We may limit ourselves, in hypothesis and conclusion, to any set of values of y, $y > 0$, denumerable or not, of which $y = 0$ is a limiting value.

Egoroff's theorem * is that if a sequence $f_n(x)$, measurable on (a, b), converges almost everywhere to $f(x)$, as n tends to infinity, the convergence is uniform except on a set of arbitrarily small measure. Accordingly if the absolute continuity of $\int f(x, y_k)\, dx$ is uniform over a sequence of positive values y_k approaching zero, and $\lim(y_k = 0+)\ f(x, y_k) = f(x, 0)$, then $f(x, 0)$ is summable, and for that sequence $\lim(y_k = 0+)\ I_1 = I_0$. It is in this sense that our theorem is a partial converse of that of Vitali.

* D-Th. Egoroff, "Sur les suites de fonctions mesurables," *Comptes Rendus Hebd. des Séances de l'Académie des Sciences*, Vol. 152 (1911), pp. 244-246.

RECIPROCAL ARRAYS AND DIOPHANTINE ANALYSIS.

By E. T. Bell.

I. The Seven Types.

1. In some applications which I have made of algebraic invariants and covariants to Diophantine analysis, the questions discussed in the present paper arise as necessary preliminaries. For reasons pointed out in § 5 these questions are of independent interest. The connection between this paper and the work on invariants, etc., will be sufficiently indicated by the two following examples.

To find all sets of integers a, b, c, d satisfying

$$(ad-bc)^2-4(ac-b^2)(bd-c^2)=0,$$

it is necessary to find all sets of non-zero integers x, y, z, u, v, w satisfying

$$xyz^2=uvw^2,$$

which is an equation of Type (V) below. When the equation of Type (V) is solved, the solution of the other presents no difficulty. Similarly for

$$abc+2fgh-af^2-bg^2-ch^2=0,$$

for which all sets of integers a, b, c, f, g, h are required. This is referred to the complete solution in non-zero integers of

$$xu_1^2=yv_1=zw_1,\ yu_2^2=zv_2=xw_2,\ zu_3^2=xv_3=yw_3,$$

which is a system of Type (VII).

2. Unless otherwise noted, *solution* shall mean all sets of non-zero values of the variables satisfying a given equation or system of equations.

A *parameter* here is a variable whose range of values is all non-zero integers. A *restricted parameter* is a variable which takes only a finite number of specified non-zero integer values (in what follows $1, -1$).

3. I shall give a straightforward, uniform method for finding the solutions of equations of each of the following types (I)-(VII). The variables are the x, y, z, u, w, all independent.

(I) $\quad x_1\cdots x_n=u_1\cdots u_n (n>1).$

(II) $\quad x_1\cdots x_n=u_1\cdots u_m (n>1, m<n).$

(III) $x_1 \cdots x_n = u_1 \cdots u_n = \cdots = w_1 \cdots w_n (n > 1),$

where the number of equal products of degree n is finite; (if the number is infinite, so are the solutions).

(IV) $x_1 \cdots x_n = u_1 \cdots u_m = \cdots = w_1 \cdots w_r (n > 1).$

Denote the power product $x_1^{a_1} \cdots x_n^{a_n}$ by $X(n)$ where $a_1, \cdots, a_n$ are constant integers > 0. With a similar notation for any power product, the next types are

(V) $X(n) = U(m),$

(VI) $X(n) = U(m) = \cdots = W(r).$

The final type is a simultaneous system of systems of type (VI), in which the power products in any row have no variable in common, but at least one power product in any row has at least one variable in common with one power product in some other row, and the system does not split into two or more independent systems.

(VII)
$$
\begin{aligned}
X_1(n_1) &= \cdots = X_r(n_r),\\
U_1(m_1) &= \cdots = U_s(m_s),\\
\cdots & \quad\quad \cdots\\
W_1(r_1) &= \cdots = W_t(r_t).
\end{aligned}
$$

Obviously (I)-(VI) can be considered as special cases of (VII). In the method followed here, however, (I) is the fundamental type. We shall next outline how repeated applications of the solution of (I) give a uniform procedure for solving (II)-(VII). For details the body of the paper must be consulted, but this summary will obviate explanations later.

4. Suppose (I) has been solved. The solution (§ 11) expresses each of x_i, u_i as a product of n independent parameters (see § 2), and in all there are n^2 such parameters in the solution. For any i, x_i, u_i in the solution have as their G. C. D. the one parameter which they have in common. Let ϕ_i be this parameter, and *denote by* $[p, q]$ *the* G. C. D. *of the non-zero integers* p, q. Then we shall refer to

$$[x_i, u_i] = \phi_i \quad (i = 1, \cdots, n)$$

as *the G. C. D. conditions on the solution,* and likewise in all similar situations. It may be emphasized once for all that the G. C. D. conditions are of the first importance in Types (V)-(VII), for without these conditions the

minimum number of parameters necessary and sufficient for the solutions can not be determined. If only formulas giving all sets of non-zero integers satisfying an equation or system of equations are desired, the G. C. D. conditions can be suppressed.

In Type (II) we first use a device of frequent application in all subsequent types, that of *making the equations formally homogeneous* by the introduction of new independent variables as factors. Rewrite (II),

$$x_1 \cdots x_n = u_1 \cdots u_m \cdots u_n, \tag{II$'$}$$

where the $n - m$ new independent variables u_j, $j > m$, are to be solved for with the rest. Apply (I) to (II′), and in the resulting solution of (II′) equate each of u_j, $j > m$, to unity. This restricts $n(n-m)$ of the n^2 parameters in the solution to range ± 1; *it will be shown that the unit factors thus introduced into the parametric expressions for* $x_1, \cdots, x_n, u_1, \cdots, u$ *can be suppressed.* Hence finally there are $n^2 - n(n-m), = nm$ parameters in the solution of (II).

Type (III) is introduced to take care of powers higher than the first in Types (V)-(VII). Write $X_n \equiv x_1 \cdots x_n$, with a similar notation for $Y_n, \cdots, W_n$. Then (III) is of the form

$$X_n = Y_n = Z_n = V_n = \cdots = T_n = U_n = W_n, \tag{III$'$}$$

which is equivalent to the "staggered" system

$$\begin{aligned} X_n &= Y_n, \\ Y_n &= Z_n, \\ Z_n &= V_n, \\ \cdots\ T_n &= U_n, \\ U_n &= W_n. \end{aligned}$$

Apply (I) to each pair in this system. From each of $Y_n, Z_n, \cdots, U_n$ we then get n equations of Type (I) in n^2 independent variables (the parameters introduced by applying (I) at the first stage), by equating, for Y_n, say, the necessarily equal values of $y_i (i = 1, \cdots, n)$. Thereafter the process is repeated for any variables (new or old parameters) for which two or more different expressions are obtained. By a definite number of applications of Type (I) the solution of (III) is obtained in terms of n^p parameters, where p is the number of equal products of n variables each in (III). Finally, by substituting back, the G. C. D. conditions on the solution are written down.

To solve (IV), proceed first as in (II), making the products formally homogeneous of degree t, where $t = \max(n, m, \cdots, r)$. Then apply (III).

In the resulting solution strike out all restricted parameters—those ranging ± 1.

The form of the solution of (IV) summarizes those of (I)-(III). There *are* $nm \cdots r$ *parameters, and each of* $x_1, \cdots, x_n$ *is a product of degree* $(nm \cdots r)/n, = m \cdots r$ *of distinct parameters; each of* $u_1, \cdots, u_m$ *is a product of degree* $(nm \cdots r)/m$ *of distinct parameters, and so on.*

With the new detail for powers of variables higher than the first, the solutions of (V)-(VII) proceed systematically as already sketched for (I)-(IV). If x^a appears as a factor, $a > 1$, it is "degraded" to $x_1 \cdots x_a$, and the new variables $x_1, \cdots, x_a$ are treated first as independent. In the resulting solution (or better, at each stage), powers are restored by solving

$$x_1 = \cdots = x_a$$

by means of (III). *The G. C. D. conditions at each stage are reduced by applying the theorem that if* $[xu, yu] = 1$, *then* $|u| = 1$. *This reduces the number of parameters by restricting those composing the factor* u *in* $|u| = 1$ *to range* ± 1. The restricted parameters are stricken out. The number of parameters may be further reduced if, say, the parameters $\phi_1, \cdots, \phi_s$ $(s > 1)$ occur in the solution only as the product $\phi_1 \cdots \phi_s$, which may then be replaced by the single parameter ϕ, thus reducing the number by $s - 1$.

Examples of all of the processes described occur in the sequel. The rapidity with which the number of parameters increases with the degree and number of equal products in a system is disconcerting but inevitable.

5. In the theory of algebraic invariants * and elsewhere it is of importance to solve completely in non-negative integers the general linear non-homogeneous Diophantine equation with positive integer coefficients. The general homogeneous (linear) equation with arbitrary integer coefficients is also required in the same connection. The algorithm of reciprocal arrays developed here for Types (I)-(VII) can be transposed to a mechanical process for finding the so-called *simple, irreducible,* or *fundamental* sets of solutions of the linear equations just mentioned, or to systems of such equations.† In fact it has been observed by Professor Morgan Ward that the multiplicative

* See, for example, Elliott, *Algebra of Quantics,* Chap. IX; Grace and Young, *Algebra of Invariants,* Chap. VI.

† The problem is also equivalent to that of finding a basis of a certain module. If to the classic theory of H. J. S. Smith for linear Diophantine systems we add the restriction that the integers in the solutions are to be non-negative, which introduces considerable difficulties, we have the *additive* dual of the general *multiplicative* problem of the present paper.

problems of this paper are abstractly identical with the additive ones for linear systems.*

II. Reciprocal Arrays.

6. We consider *pairs of square arrays A_n, A'_n, of order n, each of which contains n^2 distinct elements, and the elements in both are the same.* The ordered pair A_n, A'_n is said to be *reciprocal* if it has a certain *diagonal-row symmetry* (diagonals of A_n, rows of A'_n), which is most clearly seen from a few examples. As everything depends upon this symmetry we take space to illustrate it for $n=5, 4, 3, 2, 1$, and $n=6, 7$, from which it is obvious. The examples also show the process of *contraction*,

$$A_{n+1},\ A'_{n+1} \rightarrow A_n,\ A'_n,$$

for $n=4, 3, 2, 1$, by which a reciprocal pair of order $n+1$ is contracted to a reciprocal pair of order n, and that of *expansion*,

$$A_n,\ A'_n \rightarrow A_{n+1},\ A'_{n+1},$$

for $n=5, 6$, by which a reciprocal pair of order n is expanded to a reciprocal pair of order $n+1$. *Thus the characteristic diagonal-row symmetry of reciprocal pairs is invariant under contraction and expansion.*

In the first five examples the parts played by the accented letters are to be observed. The first array in each pair is A_n, the second A'_n.

Order 5:

$$\begin{array}{ccccc} a & b' & c & d & e \\ f & g & h' & i & j \\ k & l & m & n' & o \\ p & q & r & s & t' \\ u' & v' & w' & x' & y' \end{array} \qquad\qquad \begin{array}{ccccc} a & g & m & s & y' \\ f & l & r & x' & e \\ k & q & w' & d & j \\ p & v' & c & i & o \\ u' & b' & h' & n' & t'. \end{array}$$

Order 4; A_5, $A'_5 \rightarrow A_4$, A'_4:

$$\begin{array}{cccc} a & c' & d & e \\ f & g & i' & j \\ k & l & m & o' \\ p' & q' & r' & s' \end{array} \qquad\qquad \begin{array}{cccc} a & g & m & s' \\ f & l & r' & e \\ k & q' & d & j \\ p' & c' & i' & o'. \end{array}$$

* I first obtained the solution for Type (I) additively by finding the fundamental solution for the abstractly identical linear equation. As this is not required in the proofs here, and as Professor Ward has independently discussed the additive problem of which my example was a very simple instance, it is omitted.

Order 3; $A_4, A'_4 \rightarrow A_3, A'_3$:

$$\begin{matrix} a & d' & e \\ f & g & j' \\ k' & l' & m' \end{matrix} \qquad \begin{matrix} a & g & m' \\ f & l' & e \\ k' & d' & j'. \end{matrix}$$

Order 2; $A_3, A'_3 \rightarrow A_2, A'_2$:

$$\begin{matrix} a & e' \\ f' & g' \end{matrix} \qquad \begin{matrix} a & g' \\ f' & e'. \end{matrix}$$

Order 1; $A_2, A'_2 \rightarrow A_1, A'_1$:

$$a \qquad a$$

Any number of such contractions may be performed simultaneously by deleting any number of rows from A'_n and the corresponding diagonals from A_n. Similarly for expansion, by insertion of rows and the corresponding diagonals, with the elements in the orders indicated, as in the following examples. The capital letters are to be noticed.

Order 6; $A_5, A'_5 \rightarrow A_6, A'_6$:

in A_5, A'_5 insert rows and diagonals of asterisks as indicated next,

$$\begin{matrix} a & * & b & c & d & e \\ f & g & * & h & i & j \\ k & l & m & * & n & o \\ p & q & r & s & * & t \\ u & v & w & x & y & * \\ * & * & * & * & * & * \end{matrix} \qquad \begin{matrix} a & g & m & s & y & * \\ f & l & r & x & * & e \\ k & q & w & * & d & j \\ p & v & * & c & i & o \\ u & * & b & h & n & t \\ * & * & * & * & * & * \end{matrix}$$

which is a preliminary that can be omitted. Fill out the degenerate asterisked forms,

$$\begin{matrix} a & A & b & c & d & e \\ f & g & B & h & i & j \\ k & l & m & C & n & o \\ p & q & r & s & D & t \\ u & v & w & x & y & E \\ P & Q & R & S & T & V \end{matrix} \qquad \begin{matrix} a & g & m & s & y & V \\ f & l & r & x & T & e \\ k & q & w & S & d & j \\ p & v & R & c & i & o \\ u & Q & b & h & n & t \\ P & A & B & C & D & E \end{matrix}$$

In the next the preliminary is omitted.

Order 7; $A_6, A'_6 \rightarrow A_7, A'_7$, or $A_5, A'_5 \rightarrow A_7, A'_7$:

$$\begin{array}{ccccccc} a & A' & A & b & c & d & e \\ f & g & B' & B & h & i & j \\ k & l & m & C' & C & n & o \\ p & q & r & s & D' & D & t \\ u & v & w & x & y & E' & E \\ P & Q & R & S & T & V & F' \\ W' & P' & Q' & R' & S' & T' & V' \end{array} \qquad \begin{array}{ccccccc} a & g & m & s & y & V & V' \\ f & l & r & x & T & T' & e \\ k & q & w & S & S' & d & j \\ p & v & R & R' & c & i & o \\ u & Q & Q' & b & h & n & t \\ P & P' & A & B & C & D & E \\ W' & A' & B' & C' & D' & E' & F' \end{array}$$

If the arrays are wrapped round right circular cylinders which just take them, the rows of A'_n become the complete diagonals of A_n, which appear as unbroken arcs of helices. The diagonal-row symmetry is even more evident if the first column of A'_n be moved over to follow the last. Formal definitions follow. The genesis of the definitions is evident from any of the preceding examples.

7. The i-th row from the top of A_n is denoted by R_i, and the j-th column from the left by C_j; similarly for A'_n, R'_i, C'_j. With an obvious meaning we may write

$$A_n \equiv C_1, \cdots, C_n, \equiv \begin{array}{l} R_1 \\ \cdot \\ R_n, \end{array}$$

and similarly for A'_n.

Consider the n elements in R_i, from left to right, as forming a vector, or matrix of one row and n columns. The elements in C_j, in the order in which they occur from top to bottom, form a matrix of one column and n rows. Likewise for A'_n, C'_i, R'_j.

Now regard the transpose T_j of C_j (into a matrix of one row and n columns) as the symbol of a substitution on the n elements of C_j, and let $T_j^p\, C_j$ denote the result of applying the p-th power of T_j to the elements of C_j. Transpose the new matrix $T_j^p\, C_j$ of one row and n columns into a matrix $(T_j^p\, C_j)'$ of one column and n rows.

Then, by definition,

$$C'_j \equiv (T_j^{j-1}\, C_j)' (j=1, \cdots, n);$$

$$A_n \equiv C_1, \cdots, C_n; \; A'_n \equiv C'_1, \cdots, C'_n,$$

where A_n is any square array of n^2 distinct elements, and A_n, A'_n, in this order, are called a reciprocal pair of order n.

8. What follows is seen intuitively if the arrays or their corresponding lattices (i, j), $(i, j)'$ defined presently be imagined wrapped round cylinders as suggested in § 6.

Denote the element in row i, column j of A_n by (i, j), and similarly for A'_n, $(i, j)'$. From the definition in § 7 we then have

$$(i, 1)' = (i, 1) \quad (i = 1, \cdots, n),$$

and, if $n > 1$,

$$(i, j)' = (j + i - 1, j) \quad (j = 2, \cdots, n;\ i = 1, \cdots, n - j + 1),$$

$$(i, j)' = (j + i - n - 1, j) \quad (j = 2, \cdots, n;\ i = n - j + 2, \cdots, n).$$

From these or the definition of A'_n by powers of substitutions, all the properties of A_n, A'_n which will be required follow at once.

Let A_n be represented as part of a unit lattice on the plane of coördinates (i, j), in which the positive axis of i is drawn vertically downward, and the positive axis of j to the right. Place A_n on the unit lattice so that its first column C_1 falls along the line $j = 1$, and its first row R_1 along the line $i = 1$. The *principal diagonal* D_1 *of* A_n is

$$D_1 \equiv (1, 1), \cdots, (n, n),$$

these points occurring in the order written down the diagonal. Through each of the remaining $n - 1$ (if $n > 1$) points on C_1 draw parallels to D_1, and similarly for the remaining $n - 1$ points on R_1. We thus have $2(n - 1)$ segments lying wholly in or on A_n parallel to D_1. Read the points of A_n on these segments in the order in which they occur down from left to right.

If $n > 1$, $i > 1$, the points of A_n on the segment containing $(i, 1)$, followed by those on the segment containing $(1, n + 2 - i)$, are defined to be the *diagonal* D_i of A_n, and we have

$$(1) \qquad D_p = R'_p \ (p = 1, \cdots, n).$$

Now place A'_n on a unit lattice in exactly the same way, and consider its *principal right-to-left diagonal* D'_n;

$$D'_n : (1, n)', (2, n - 1)', \cdots, (n, 1)'.$$

Then we see that

(2) *The elements of* A'_n *on* D'_n *are the elements of* A_n *on* R_n *in reverse order (right to left on* R_n*).*

From (1), (2) follows the fundamental property of the reciprocal pair A_n, A'_n, which is illustrated by the contractions and expansions in § 6. From A_n delete D_n and R_n; from A'_n delete D'_n and R'_n. Leave the first column in the array thus obtained from A_n unchanged, and shift each element after the deleted one in every row (except the last, in which no elements remain),

one place to the left. In the array obtained from A'_n, shift each element before the deleted one in every row one place to the right. Then *the resulting square arrays* A_{n-1}, A'_{n-1} *have* $(n-1)^2$ *elements each and are a reciprocal pair.* We say that A_{n-1}, A'_{n-1} have been derived from A_n, A'_n by *contraction,* and write

$$A_n, A'_n \rightarrow A_{n-1}, A'_{n-1}.$$

The inverse process of *expansion,*

$$A_n, A'_n \rightarrow A_{n+1}, A'_{n+1},$$

is carried out thus: the new R'_{n+1} of $n+1$ new elements is adjoined to A'_n below R'_n, and is inserted as the new D_{n+1} in A_n; the new R_{n+1} is adjoined to A_n and is inserted as the new D'_{n+1}.

Contraction and expansion generate reciprocal pairs from reciprocal pairs.

9. The *normal form* of the reciprocal pair A_n, A'_n of order n is merely a matter of convenience in notation, and is as follows: The n^2 elements of A_n are the integers $1, \cdots, n^2$ arranged in a square array, in which the element (i, j) in row i, column j, is the j-th term of the arithmetical progression whose first term is 1 and whose common difference is n,

$$(i, j) = i + (j-1)n.$$

From this A_n the normal A'_n in the reciprocal pair A_n, A'_n is written down by means of substitutions, as in the definition (§ 7) of reciprocal pairs, or more simply by a rule which is obvious from the normal pair for $n = 6$.

1,	7,	13,	19,	25,	31	1,	8,	15,	22,	29,	36
2,	8,	14,	20,	26,	32	2,	9,	16,	23,	30,	31
3,	9,	15,	21,	27,	33	3,	10,	17,	24,	25,	32
4,	10,	16,	22,	28,	34	4,	11,	18,	19,	26,	33
5,	11,	17,	23,	29,	35	5,	12,	13,	20,	27,	34
6,	12,	18,	24,	30,	36	6,	7,	14,	21,	28,	35.

10. One further consequence, the possibility of *absorption of units,* must be considered for the reciprocal pair A_n, A'_n, $n > 1$.

The effect of deleting r rows, $r < n$, of A'_n is to delete also r diagonals of A_n. After compression to fill the r vacancies in each row of A_n, let A_n become $A_{n,n-r}$, a rectangular array of n rows and $n-r$ columns. A'_n becomes $A'_{n-r,n}$ of $n-r$ rows and n columns. By renumbering, if necessary, the r deleted rows of A'_n may be taken as the last r. Imagine the elements of the r deleted diagonals of A_n repalced by asterisks, and advance all diagonals a

sufficient number of places to bring them into coincidence with the last. Finally, pair the single asterisk now in each row with the element immediately following it in the row (the first follows the last). The application of this to Diophantine analysis is made from the following interpretation of the operations described.

Let A_n, A'_n be in normal form (§ 9), and let $1, \cdots, n^2$ be the suffixes of n^2 independent parameters (§ 2) ϕ_k. Denote the product of the n parameters ϕ_k whose suffixes are in R_i by x_i, and similarly for R'_i and u_i. Restrict the nr parameters in u_p $(p = n, n-1, \cdots, n-r+1)$ by the r conditions $u_p = 1$. Then each of the nr takes one of the values $1, -1$, and none or an even number of those in a particular u_p take the value -1. The nr positive or negative units in any permissible choice can be absorbed as factors in the x_i in such a way that the product of units and a parameter can be replaced by a new parameter (ranging all non-zero integers, § 2), and *the new x_i range the same sets of values, except possibly for order, as the old x_i. The remaining u's are unaffected.* Hence, the units introduced as described may be stricken out without affecting the ranges of the modified x's and the remaining u's.

III. Types (I)-(IV)

11. The ϕ's denote parameters (§ 2); for the meaning of *solution* see § 2.

The solution of

$$x_1 \cdots x_n = u_1 \cdots u_n (n > 1)$$

is

$$x_i = \phi_{(i,1)} \cdots \phi_{(i,n)},\quad u_i = \phi_{(i,1)'} \cdots \phi_{(i,n)'} (i = 1, \cdots, n),$$

where (i, j), $(i, j)'$ are the elements in row i, column j of the reciprocal arrays A_n, A'_n, respectively (§ 7), *and the n^2 parameters are independent, subject to the* G. C. D. *conditions* (§ 4)

$$[x_i, u_i] = \phi_{(i,1)} (= \phi_{(i,1)'}) (i = 1, \cdots, n).$$

For applications to subsequent types it is advantageous to take the G. C. D. conditions in the equivalent form

$$1 = [x_i/\phi_{(i,1)},\ u_i/\phi_{(i,1)'}] \ (i = 1, \cdots, n).$$

If A_n, A'_n are in normal form, the solution is also said to be in normal form. As an example we write down the normal form of the solution of

$$x_1 \cdots x_6 = u_1 \cdots u_6:$$

$$x_1 = \phi_1\phi_7\ \phi_{13}\phi_{19}\phi_{25}\phi_{31},\quad u_1 = \phi_1\phi_8\ \phi_{15}\phi_{22}\phi_{29}\phi_{36},$$

$$x_2 = \phi_2\phi_8\ \phi_{14}\phi_{20}\phi_{26}\phi_{32},\quad u_2 = \phi_2\phi_9\ \phi_{16}\phi_{23}\phi_{30}\phi_{31},$$

$$x_3 = \phi_3\phi_9\,\phi_{15}\phi_{21}\phi_{27}\phi_{33},\quad u_3 = \phi_3\phi_{10}\phi_{17}\phi_{24}\phi_{25}\phi_{32},$$
$$x_4 = \phi_4\phi_{10}\phi_{16}\phi_{22}\phi_{28}\phi_{34},\quad u_4 = \phi_4\phi_{11}\phi_{18}\phi_{19}\phi_{26}\phi_{33},$$
$$x_5 = \phi_5\phi_{11}\phi_{17}\phi_{23}\phi_{29}\phi_{35},\quad u_5 = \phi_5\phi_{12}\phi_{13}\phi_{20}\phi_{27}\phi_{34},$$
$$x_6 = \phi_6\phi_{12}\phi_{18}\phi_{24}\phi_{30}\phi_{36},\quad u_6 = \phi_6\phi_7\,\phi_{14}\phi_{21}\phi_{28}\phi_{35};$$
$$1 = [x_i/\phi_i,\ u_i/\phi_i]\ (i = 1, \cdots, 6).$$

12. To prove the result in § 11 we show first and independently that it holds for $n = 2$, 3 and then complete the proof by mathematical induction.

If x, y are non-zero integers, $x \mid y$ is read "x divides y" (arithmetically). Hence, if x, y, z are non-zero integers, $x \mid y$ and $y = xz$ are equivalent statements. As before, $[x, y]$ denotes the G. C. D. of x, y. Let

$$x_1x_2 = u_1u_2;\ [x_1, u_1] = \delta;\ x_1 = \delta x'_1,\ u_1 = \delta u'_1.$$

Then

$$x'_1x_2 = u'_1u_2,\ [x'_1, u'_1] = 1,$$

and therefore

$$x'_1 \mid u_2,\ u'_1 \mid x_2;\ \ u_2 = \lambda x'_1,\ x_2 = \mu u'_1;$$
$$x'_1\mu u'_1 = u'_1\lambda x'_1;\ \lambda = \mu;$$
$$x_1 = \delta x'_1,\ x_2 = \mu u'_1,\ u_1 = \delta u'_1,\ u_2 = \mu x'_1,$$
$$[x_1, u_1] = \delta[x'_1, u'_1] = \delta,\ [x_2, u_2] = \mu[u'_1, x'_1] = \mu.$$

The change of notation

$$\delta,\ x'_1,\ \mu,\ u'_1 = \phi_1,\ \phi_3,\ \phi_2,\ \phi_4$$

completes the proof for $n = 2$. The similar proof for $n = 3$ is given in connection with the work mentioned in § 1, and may be omitted.

The induction from n to $n + 1$ is most clearly seen by following it through with an example, which illustrates all the features of the general case unencumbered by notation. Take $n = 5$, and assume that the result in § 11 holds for 2, 3, 5 (general, 2, 3, n). We shall prove it true for $n = 6$. Apply the theorem for $n = 2$ to

$$(x_1 \cdots x_5)x_6 = (u_1 \cdots u_5)u_6.$$

Then

$$x_1 \cdots x_5 = \delta\phi,\ u_1 \cdots u_5 = \delta\psi,$$
$$x_6 = \mu\psi,\ u_6 = \mu\phi,\ [\phi, \psi] = 1.$$

Fill out the first pair to make them formally homogeneous of degree 5 (in general case, n),

$$x_1 \cdots x_5 = \delta\phi a_3a_4a_5,\ u_1 \cdots u_5 = \delta\psi b_3b_4b_5,$$
$$(1 = a_3a_4a_5 = b_3b_4b_5).$$

By hypothesis, § 11 holds for each of these equations of degree $n = 5$. Hence

$$\begin{array}{ll} x_1 = \phi_1\phi_6\ \phi_{11}\phi_{16}\phi_{21}, & \delta\ = \phi_1\phi_7\ \phi_{13}\phi_{19}\phi_{25}, \\ x_2 = \phi_2\phi_7\ \phi_{12}\phi_{17}\phi_{22}, & \phi\ = \phi_2\phi_8\ \phi_{14}\phi_{20}\phi_{21}, \\ x_3 = \phi_3\phi_8\ \phi_{13}\phi_{18}\phi_{23}, & a_3 = \phi_3\phi_9\ \phi_{15}\phi_{16}\phi_{22}, \\ x_4 = \phi_4\phi_9\ \phi_{14}\phi_{19}\phi_{24}, & a_4 = \phi_4\phi_{10}\phi_{11}\phi_{17}\phi_{23}, \\ x_5 = \phi_5\phi_{10}\phi_{15}\phi_{20}\phi_{25}, & a_5 = \phi_5\phi_6\ \phi_{12}\phi_{18}\phi_{24}; \end{array}$$

$$[x_1, \delta] = \phi_1,\ [x_2, \phi] = \phi_2,\ [x_j, a_j] = \phi_j (j = 3, \cdots, 5).$$

In the general case, $j = 3, \cdots, n$. The solution of $u_1 \cdots u_5 = \delta\psi b_3 b_4 b_5$ is similarly written down in terms of parameters $\psi_1, \cdots, \psi_{25}$. In the above solution delete or absorb the units introduced by $1 = a_3 a_4 a_5$ (general, $1 = a_3 \cdots a_n$), and mark the places of the units by asterisks. This leaves (general) D_1, D_2 in the above expressions for the x's. Similarly for the u's. Then we have

$$\begin{array}{llllll} x_1 = \phi_1 & * & * & * & \phi_{21}, & u_1 = \psi_1 \ * \ * \ * \ \psi_{21}, \\ x_2 = \phi_2 & \phi_7 & * & * & *\ , & u_2 = \psi_2 \ \psi_7 \ * \ * \ *\ , \\ x_3 = * & \phi_8 & \phi_{13} & * & *\ , & u_3 = * \ \psi_8 \ \psi_{13} \ * \ *\ , \\ x_4 = * & * & \phi_{14} & \phi_{19} & *\ , & u_4 = * \ * \ \psi_{14} \ \psi_{19} \ *\ , \\ x_5 = * & * & * & \phi_{20} & \phi_{25}, & u_5 = * \ * \ * \ \psi_{20} \ \psi_{25}, \end{array}$$

Equate (general) the values of δ,

$$\phi_1\phi_7\phi_{13}\phi_{19}\phi_{25} = \psi_1\psi_7\psi_{13}\psi_{19}\psi_{25},$$

the respective products being (general) the principal diagonals, respectively, of the starred arrays. Hence the equation is of degree $5 (= n)$, and we may again apply § 11:

$$\begin{array}{ll} \phi_1\ = \theta_1\theta_6\ \theta_{11}\theta_{16}\theta_{21}, & \psi_1\ = \theta_1\theta_7\ \theta_{13}\theta_{19}\theta_{25}, \\ \phi_7\ = \theta_2\theta_7\ \theta_{12}\theta_{17}\theta_{22}, & \psi_7\ = \theta_2\theta_8\ \theta_{14}\theta_{20}\theta_{21}, \\ \phi_{13} = \theta_3\theta_8\ \theta_{13}\theta_{18}\theta_{23}, & \psi_{13} = \theta_3\theta_9\ \theta_{15}\theta_{16}\theta_{22}, \\ \phi_{19} = \theta_4\theta_9\ \theta_{14}\theta_{19}\theta_{24}, & \psi_{19} = \theta_4\theta_{10}\theta_{11}\theta_{17}\theta_{23}, \\ \phi_{25} = \theta_5\theta_{10}\theta_{15}\theta_{20}\theta_{25}, & \psi_{25} = \theta_5\theta_6\ \theta_{12}\theta_{18}\theta_{24}; \end{array}$$

$$[\phi_1, \psi_1] = \theta_1,\ [\phi_7, \psi_7] = \theta_2,\ [\phi_{13}, \psi_{13}] = \theta_3,\ [\phi_{19}, \psi_{19}] = \theta_4,\ [\phi_{25}, \psi_{25}] = \theta_5.$$

We now expand this reciprocal pair, and hence obtain a reciprocal pair of order 6 (general, $n + 1$), using for the new row and diagonal of the first the elements in the above expressions for x_6, u_6 (general x_{n+1}, u_{n+1}), respectively,

$$x_6 = \mu\psi_2\psi_8\psi_{14}\psi_{20}\psi_{21},\ u_6 = \mu\phi_2\phi_8\phi_{14}\phi_{20}\phi_{21},$$

after permuting the ψ's, ϕ's (general) *into the orders in*

$$x_6 = \mu\psi_{20}\psi_{14}\psi_8\psi_2\psi_{21},\ u_6 = \mu\phi_{21}\phi_2\phi_8\phi_{14}\phi_{20}.$$

(General: the last ψ is left in place; the others are written in reverse order;

the ϕ's are cyclically reversed). Compare now the θ-arrays and the starred pair. Then (general), since, by the binary solution at the beginning, we have $[\phi, \psi] = 1$, it follows that

$$[x_i, u_i] = \theta_i \ (i = 1, \cdots, n), \ [x_{n+1}, u_{n+1}] = \mu;$$

here $n = 5$.

Finally note the expressions for x_i, u_i $(i = 1, \cdots, n)$ in the starred pair, and compare with the expanded θ-pair. This completes the induction. For $n = 5$ the last step, the expansion, is

$$\begin{aligned}
x_1 &= \theta_1\phi_{21}\theta_6 \ \theta_{11}\theta_{16}\theta_{21}, & u_1 &= \theta_1\theta_7 \ \theta_{13}\theta_{19}\theta_{25}\psi_{21}, \\
x_2 &= \theta_2\theta_7 \ \phi_2 \ \theta_{12}\theta_{17}\theta_{22}, & u_2 &= \theta_2\theta_8 \ \theta_{14}\theta_{20}\psi_2 \ \theta_{21}, \\
x_3 &= \theta_3\theta_8 \ \theta_{13}\phi_8 \ \theta_{18}\theta_{23}, & u_3 &= \theta_3\theta_9 \ \theta_{15}\psi_8 \ \theta_{16}\theta_{22}, \\
x_4 &= \theta_4\theta_9 \ \theta_{14}\theta_{19}\phi_{14}\theta_{24}, & u_4 &= \theta_4\theta_{10}\psi_{14}\theta_{11}\theta_{17}\theta_{23}, \\
x_5 &= \theta_5\theta_{10}\theta_{15}\theta_{20}\theta_{25}\phi_{20}, & u_5 &= \theta_5\psi_{20}\theta_6 \ \theta_{12}\theta_{18}\theta_{24}, \\
x_6 &= \mu \ \psi_{20}\psi_{14}\psi_8\psi_2 \ \psi_{21}, & u_6 &= \mu \ \phi_{21}\phi_2 \ \phi_8 \ \phi_{14}\phi_{20};
\end{aligned}$$

$$[x_i, u_i] = \theta_i (i = 1, \cdots, 5), \ [x_6, u_6] = \mu.$$

By relettering and renumbering of parameters, a frequently useful device, we throw this into the normal form, as in the example in § 11.

14. *The solution of*

$$x_1 \cdots x_n = u_1 \cdots u_m (n > 1, m < n)$$

is written down from § 11 *by deleting the last* $n - m$ *u's in the solution there given, and striking out of the x's those diagonals corresponding* [*by the diagonal-row symmetry of* A_n, A'_n, § 8 (1)] *to the deleted u's. In the result the G. C. D. conditions in* § 11 *are replaced by what they become when the deleted parameters are replaced by units.* Proof by absorption of units (§ 10).

If in a particular $1 = [x, y]$, either x or y is a unit, the condition is superfluous, and is suppressed. The G. C. D. conditions become

$$[x_i, u_i] = p_i (i = 1, \cdots m),$$

where p_i is the *algebraic* H. C. F. of x_i, u_i.

As an example we write down from the example in § 11 the solution of

$$x_1x_2x_3x_4x_5x_6 = u_1u_2.$$

The reduced arrays are

1, 31	1, 8, 15, 22, 29, 36
2, 8	2, 9, 16, 23, 30,31.
3, 15	
16, 22	
23, 29	
30, 36	

Hence, after suitable renumbering, we have

$$\begin{aligned} x_1 &= \phi_1\phi_7\,, & u_1 &= \phi_1\phi_8\phi_9\phi_{10}\phi_{11}\phi_{12}, \\ x_2 &= \phi_2\phi_8\,, & u_2 &= \phi_2\phi_3\phi_4\phi_5\,\phi_6\,\phi_7\,; \\ x_3 &= \phi_3\phi_9\,, \\ x_4 &= \phi_4\phi_{10}, \\ x_5 &= \phi_5\phi_{11}, \\ x_6 &= \phi_6\phi_{12}, \end{aligned}$$

$$[x_1, u_1] = \phi_1,\ [x_2, u_2] = \phi_2.$$

The solution for Types III, IV is sufficiently evident from §§ 4, 11-13. Proceeding as sketched in § 4, we find that Type III (§ 4) demands n^p parameters, and from this solution, by absorption of units, we reach the conclusion italicized in § 4 for Type IV.

At each stage the G. C. D. *conditions appropriate to it* (given by applying Type I) *are included.** This precaution, possibly unnecessary as remarked in the footnote, is inserted to take care of every possible reduction of the number of parameters in Types V-VII.

The form in which the G. C. D. conditions come out *in the last set* is interesting. We may arrange any system of Type IV in the form

$$x_1 \cdots x_n = y_1 \cdots y_m = z_1 \cdots z_r = \cdots = w_1 \cdots w_s,$$
$$1 < n \geqq m \geqq r \geqq \cdots \geqq s.$$

Consider the parametric expressions for the variables from *consecutive pairs* of products as given in the solution. Take any pair, say the second, $y_1 \cdots y_m = z_1 \cdots z_r$. This contributes to the total last set of G. C. D. conditions

$$[y_i, z_i] = p_i (i = 1, \cdots, r),$$

where p_i is the algebraic H. C. F. of y_i, z_i when expressed parametrically.

As the example in the next part incidentally illustrates this section, we pass to the remaining types.

IV. Types (V)-(VII)

15. With what has been given in § 4 it will suffice to show the working for

* I have not considered whether the solution obtained by ignoring all G. C. D. conditions except the last set given by the method is more redundant than that in which *all* G. C. D. conditions arising in the course of the solution are retained. It is obvious that all sets of non-zero integers satisfying the system are given if only the last set is retained, but it is not proved that this solution admits redundancies excluded by imposing all the sets of conditions. This question is of importance in the additive isomorph (§ 5).

$$x_1x_2^2 = y_1y_2 = z_1z_2, \tag{1}$$

which is one stage of the system mentioned in § 1, second example.

The first step reduces (1) formally to Type III,

$$x_1x_2x_3 = y_1y_2y_3 = z_1z_2z_3, \tag{2}$$

by making all products of the same degree and degrading powers. Having solved this we absorb the units introduced into (1) by setting $y_3 = 1$, $z_3 = 1$, and deleting the corresponding parameters in the solution of (2). Thus we have the solution of

$$x_1x_2x_3 = y_1y_2 = z_1z_2. \tag{3}$$

In this solution we equate the parametric expressions for x_2, x_3, apply Type I to solve the resulting equation for all the parameters concerned, and thus reach the "crude" solution of (1). A crude solution is one in which G. C. D. conditions are neglected. This is the straightforward mechanical way of proceeding, to defer G. C. D. conditions till the last step. It is a great saving of labor, however, to attend to them at every step, as the number of parameters is thus reduced as rapidly as possible, and subsequent steps are correspondingly simplified.

To solve (2), stagger it, and apply § 11 to write down the solutions of

$$x_1x_2x_3 = y_1y_2y_3,\ y_1y_2y_3 = z_1z_2z_3:$$

$$x_1 = \phi_1\phi_4\phi_7,\ y_1 = \phi_1\phi_5\phi_9 = \psi_1\psi_4\psi_7,\ z_1 = \psi_1\psi_5\psi_9,$$

$$x_2 = \phi_2\phi_5\phi_8,\ y_2 = \phi_2\phi_6\phi_7 = \psi_2\psi_5\psi_8,\ z_2 = \psi_2\psi_6\psi_7,$$

$$x_3 = \phi_3\phi_6\phi_9,\ y_3 = \phi_3\phi_4\phi_8 = \psi_3\psi_6\psi_9,\ z_3 = \psi_3\psi_4\psi_8;$$

$$[x_i, y_i] = \phi_i,\ [y_i, z_i] = \psi_i\,(i = 1, 2, 3).$$

Apply § 11 to equal values of y_i. Let the parameters for the equations from y_1, y_2, y_3 be α's, β's, γ's, respectively. The solutions can be written down mechanically, so we give only the results of substituting them into the preceding formulas to give the final forms for x_i, y_i, z_i.

$$x_1 = \alpha_1\alpha_4\alpha_7\gamma_2\gamma_5\gamma_8\beta_3\beta_6\beta_9,$$

$$x_2 = \beta_1\beta_4\beta_7\alpha_2\alpha_5\alpha_8\gamma_3\gamma_6\gamma_9,$$

$$x_3 = \gamma_1\gamma_4\gamma_7\beta_2\beta_5\beta_8\alpha_3\alpha_6\alpha_9;$$

$$y_1 = \alpha_1\alpha_4\alpha_7\alpha_2\alpha_5\alpha_8\alpha_3\alpha_6\alpha_9,$$

$$y_2 = \beta_1\beta_4\beta_7\beta_2\beta_5\beta_8\beta_3\beta_6\beta_9,$$

$$y_3 = \gamma_1\gamma_4\gamma_7\gamma_2\gamma_5\gamma_8\gamma_3\gamma_6\gamma_9;$$

$$z_1 = \alpha_1\alpha_5\alpha_9\beta_2\beta_6\beta_7\gamma_3\gamma_4\gamma_8,$$

$$z_2 = \beta_1\beta_5\beta_9\gamma_2\gamma_6\gamma_7\alpha_3\alpha_4\alpha_8,$$

$$z_3 = \gamma_1\gamma_5\gamma_9\alpha_2\alpha_6\alpha_7\beta_3\beta_4\beta_8;$$

with the G. C. D. conditions

$$[x_1, y_1] = \alpha_1\alpha_4\alpha_7,\ [x_2, y_2] = \beta_1\beta_4\beta_7,\ [x_3, y_3] = \gamma_1\gamma_4\gamma_7,$$
$$[y_1, z_1] = \alpha_1\alpha_5\alpha_9,\ [y_2, z_2] = \beta_1\beta_5\beta_9,\ [y_3, z_3] = \gamma_1\gamma_5\gamma_9,$$

and the preceding set, obtained at the second stage,

$$1 = [\alpha_4\alpha_7,\ \alpha_5\alpha_9] = [\alpha_5\alpha_8,\ \alpha_6\alpha_7] = [\alpha_6\alpha_9,\ \alpha_4\alpha_8],$$
$$1 = [\beta_4\beta_7,\ \beta_5\beta_9] = [\beta_5\beta_8,\ \beta_6\beta_7] = [\beta_6\beta_9,\ \beta_4\beta_8],$$
$$1 = [\gamma_4\gamma_7,\ \gamma_5\gamma_9] = [\gamma_5\gamma_8,\ \gamma_6\gamma_7] = [\gamma_6\gamma_9,\ \gamma_4\gamma_8],$$

where the alternative form noted in § 11 has been used, as it is shorter here. This completes the solution of (2). The entire solution could have been written down mechanically; notice the array of parameters and its composition out of reciprocal pairs of order 3.

To solve (3), we set $y_3 = 1$, $z_3 = 1$ in the preceding, and delete the parameters concerned. We shall not renumber nor permute the remaining parameters. The solution of (3) is

$$x_1 = \alpha_1\alpha_4\beta_6\beta_9,$$
$$x_2 = \beta_1\beta_7\alpha_5\alpha_8,$$
$$x_3 = \beta_2\beta_5\alpha_3\alpha_9\,;$$

$$y_1 = \alpha_1\alpha_4\alpha_5\alpha_8\alpha_3\alpha_9,\quad z_1 = \alpha_1\alpha_5\alpha_9\beta_2\beta_6\beta_7,$$
$$y_2 = \beta_1\beta_7\beta_2\beta_5\beta_6\beta_9,\quad z_2 = \beta_1\beta_5\beta_9\alpha_3\alpha_4\alpha_8\,;$$
$$1 = [\beta_6\beta_9,\ \alpha_5\alpha_8\alpha_3\alpha_9] = [\alpha_5\alpha_8,\ \beta_2\beta_5\beta_6\beta_9] = [\alpha_3\alpha_4\alpha_8,\ \beta_2\beta_6\beta_7]\,;$$
$$1 = [\alpha_4,\ \alpha_5\alpha_9] = [\alpha_9,\ \alpha_4\alpha_8] = [\beta_7,\ \beta_5\beta_9] = [\beta_5,\ \beta_6\beta_7].$$

The missing conditions are accounted for by degenerations to the form $[x, 1] = 1$.

From this solution we get that of (1) by setting $x_2 = x_3$:

$$\beta_1\beta_7\alpha_5\alpha_8 = \beta_2\beta_5\alpha_3\alpha_9.$$

It is here that the G. C. D. conditions play their part. By § 11 we write down

$$\beta_1 = \theta_1\theta_5\theta_9\,\theta_{13},\quad \beta_2 = \theta_1\theta_6\theta_{11}\theta_{16},$$
$$\beta_7 = \theta_2\theta_6\theta_{10}\theta_{14},\quad \beta_5 = \theta_2\theta_7\theta_{12}\theta_{13},$$

$$\alpha_5 = \theta_3\theta_7\theta_{11}\theta_{15},\quad \alpha_3 = \theta_3\theta_8\theta_9\,\theta_{14},$$
$$\alpha_8 = \theta_4\theta_8\theta_{12}\theta_{16},\quad \alpha_9 = \theta_4\theta_5\theta_{10}\theta_{15}\,;$$
$$[\beta_1, \beta_2] = \theta_1,\ [\beta_7, \beta_5] = \theta_2,\ [\alpha_5, \alpha_3] = \theta_3,\ [\alpha_8, \alpha_9] = \theta_4.$$

Substitute for $\beta_1, \beta_7, \alpha_5, \alpha_8, \beta_2, \beta_5, \alpha_3, \alpha_9$ in the G. C. D. conditions of the preceding step (the solution of (3)). At this step, *retain only those conditions of the form* $[x, y] = 1$ *in which each of* x, y *contains as an algebraic factor at least one of* $\beta_1, \beta_7, \alpha_5, \alpha_8, \beta_2, \beta_5, \alpha_3, \alpha_9$. The algebraic H. C. F.'s of the x, y in such $[x, y]$ are read off by inspection of the above solution,

and give the θ's which are to be deleted. In any such $[x, y]$, only those parameters in x, y of the solution of (3) need be retained which are among $\beta_1, \beta_7, \alpha_5, \alpha_8, \beta_2, \beta_5, \alpha_3, \alpha_9$. Here we get

$$1 = [\alpha_5\alpha_8, \beta_2\beta_5] = [\alpha_3\ \alpha_8, \beta_2\beta_7] = [\alpha_9, \alpha_8] = [\beta_7, \beta_5].$$

Hence

$$\theta_7, \theta_{11}, \theta_{12}, \theta_{16};\ \theta_{14};\ \theta_4;\ \theta_2$$

are to be deleted, and we have

$$\begin{aligned} \beta_1 &= \theta_1\theta_5\theta_9\theta_{13}, & \beta_2 &= \theta_1\theta_6, \\ \beta_7 &= \theta_6\theta_{10}, & \beta_5 &= \theta_{13}, \\ \alpha_5 &= \theta_3\theta_{15}, & \alpha_3 &= \theta_3\theta_8\theta_9, \\ \alpha_8 &= \theta_8, & \alpha_9 &= \theta_5\theta_{10}\theta_{15}, \end{aligned}$$

with the corresponding reduced G. C. D. conditions. From this we get the solution of (1), after renumbering the parameters,

$$\begin{aligned} &\alpha_1, \alpha_4, \beta_6, \beta_9;\ \theta_1, \theta_3, \theta_5, \theta_6, \theta_8, \theta_9, \theta_{10}, \theta_{13}, \theta_{15}, \\ &\qquad \rightarrow \phi_1, \phi_2, \phi_3, \phi_4;\ \psi_1, \psi_2, \psi_3, \psi_4, \psi_5, \psi_6, \psi_7, \psi_8, \psi_9. \end{aligned}$$

The solution of (1) is

$$\begin{aligned} x_1 &= \phi_1\phi_2\phi_3\phi_4, \\ x_2 &= \psi_1\psi_2\psi_3\psi_4\psi_5\psi_6\psi_7\psi_8\psi_9; \\ y_1 &= \phi_1\phi_2(\psi_2\psi_5\psi_9)^2\psi_3\psi_6\psi_7, \\ y_2 &= \phi_3\phi_4(\psi_1\psi_4\psi_8)^2\psi_3\psi_6\psi_7; \\ z_1 &= \phi_1\phi_3(\psi_4\psi_7\psi_9)^2\psi_1\psi_2\psi_3, \\ z_2 &= \phi_2\phi_4(\psi_5\psi_6\psi_8)^2\psi_1\psi_2\psi_3, \end{aligned}$$

with the G. C. D. conditions that each of the following is 1:

$$\begin{aligned} &[\phi_3\phi_4, \psi_2\psi_3\psi_5\psi_6\psi_7\psi_9], [\psi_2\psi_5\psi_9, \phi_3\phi_4\psi_1\psi_4\psi_8], \\ &[\phi_2\psi_2\psi_5\psi_6, \phi_3\psi_1\psi_4\psi_7]; \\ &[\phi_2, \psi_2\psi_3\psi_7\psi_9], [\psi_2\psi_7\psi_9, \phi_2\psi_5], [\psi_4\psi_7, \phi_4\psi_8]. \\ &[\psi_8, \phi_3\psi_4\psi_7]; \\ &[\psi_3\psi_6\psi_8, \psi_4], [\psi_9, \psi_5\psi_6]. \end{aligned}$$

The total number of parameters in the solution is thus 13, and this agrees with the number given by Professor Ward's general formula. It is interesting to notice that the 13 comes out in his formula as $2^2 + 3^2$, as is accounted for by the form of (1).

The G. C. D. conditions in any type may be dropped; the modified solution will also give all sets of non-zero integers satisfying the system, but with avoidable duplications. For applications to further systems, all G. C. D. conditions must be retained, if the final solution is to be in the least number of parameters.

A TYPE OF MULTIPLICATIVE DIOPHANTINE SYSTEM.

By Morgan Ward.

1. Consider the system of M equations in the $K+L$ unknowns $x_1, \cdots, x_K, y_1, \cdots, y_L$

$$\text{(S)} \qquad A_i x_1^{a_{i1}} x_2^{a_{i2}} \cdots x_K^{a_{iK}} = B_i y_1^{b_{i1}} y_2^{b_{i2}} \cdots y_L^{b_{iL}}, \qquad (i = 1, 2, \cdots, M).$$

The exponents a, b are assumed to be positive integers or zero, while the constants A, B are positive integers.

The problem of determining all the real positive * solutions of (S) is a trivial one; for if we let

$$z_1 = \log x_1, \cdots, z_K = \log x_K, \ w_1 = \log y_1, \cdots, w_L = \log y_L, \ e_i = \log(A_i/B_i), \quad (i = 1, \cdots, M),$$

then on taking the logarithm of both sides of each equation in (S) we obtain the linear system

$$\text{(E)} \quad a_{i1}z_1 + \cdots + a_{iK}z_K - b_{i1}w_1 - \cdots - b_{iL}w_L = e_i, \quad (i = 1, \cdots, M).$$

The solution of (S) is thus effectively reduced to a mere inspection of the matrix of the coefficients of (E).

On the other hand, the problem of determining all positive integral solutions of (S) is distinctly non-trivial, and offers several interesting and unexpected features.† To give an idea of the difficulties involved, if we seek to replace (S) by the linear system (E), we must add the restrictions that $z_1, \cdots, w_L$ be non-negative, and that $e^{z_1}, \cdots, e^{w_L}$ be rational integers. But to select from the totality of solutions of (E) the particular solutions which meet these restrictions appears to be as difficult as to solve the original system (S).

For a direct attack upon this problem, the reader may consult the paper of Bell's already referred to. The method I develop here is indirect. It is, however, strictly arithmetical, being based upon the fundamental theorem of rational arithmetic—unique decomposition into prime factors. It accordingly would not be applicable if we were attempting to find all solutions of (S) in an arbitrary domain of integrity.‡

* The negative solutions may be immediately obtained from the positive on considering the parity of the a and b.

† E. T. Bell, "Reciprocal arrays and diophantine analysis," this Journal, Vol. 55 (1933), pp. 50-66. In this paper a general non-tentative method for solving the system (M) is developed.

‡ van der Waerden, *Algebra*, Part I, Berlin (1931), p. 39.

The essentials of the method are as follows. We consider along with (S) a more special system (M) obtained on setting all the constants A and B equal to unity:

$$\text{(M)} \qquad x_1^{a_{i1}}x_2^{a_{i2}}\cdots x_K^{a_{iK}} = y_1^{b_{i1}}y_2^{b_{i2}}\cdots y_L^{b_{iL}}, \qquad (i = 1, 2, \cdots, M).$$

We then show that there exists a correspondence between the solutions of (M) in positive integers x and y and the solutions of the linear system

$$\text{(A)} \qquad a_{i1}z_1 + a_{i2}z_2 + \cdots + a_{iK}z_K = b_{i1}w_1 + b_{i2}w_2 + \cdots + b_{iL}w_L, \qquad (i = 1, 2, \cdots, M),$$

in non-negative integers z and w. This correspondence is of a dual character, so that any theorem about the solutions of (A) yields a theorem about the solutions of (M) and vice-versa. Since the broad outlines of the theory of the solution of (A) are known,* we obtain without effort considerable information about the solutions of (M). A slight extension of the method allows us finally to discuss the general system (S).

2. We must first lay down a few definitions. The systems (M) and (A) will be said to be *associated.* By a solution of (S) or (M) we shall mean a solution in positive integers, and by a solution of (A) a solution in non-negative integers. To avoid trivialities, we shall furthermore assume that (S) actually has solutions.

We shall find it convenient to represent a solution $\xi_1, \xi_2, \cdots, \xi_K, \eta_1, \eta_2, \cdots, \eta_L$ of any one of the three systems (S), (M) or (A) under consideration as a one-rowed matrix,†

$$[\xi;\eta] = [\xi_1, \xi_2, \cdots, \xi_K, \eta_1, \eta_2, \cdots, \eta_L].$$

If

$$[\xi';\eta'] = [\xi'_1, \xi'_2, \cdots, \xi'_K, \eta'_1, \eta'_2, \cdots, \eta'_L]$$

is a second such solution, then the matrix

$$[\xi + \xi';\eta + \eta'] = [\xi_1 + \xi'_1, \cdots, \eta_L + \eta'_L]$$

is called the sum of the solutions $[\xi;\eta]$, $[\xi';\eta']$ and expressed as usual by the notation

$$[\xi + \xi';\eta + \eta'] = [\xi;\eta] + [\xi';\eta'].$$

In like manner, the product of two solutions is expressed by

$$[\xi\xi';\eta\eta'] = [\xi;\eta]\cdot[\xi';\eta'],$$

* See for example, Grace and Young, *Algebra of Invariants*, Cambridge (1903), pp. 102-106.

† Bell, *Algebraic Arithmetic*, pp. 15-16.

and the identity of any two solutions by

$$[\xi;\eta]=[\xi';\eta'].$$

Finally, if t is any integer,

$$t[\xi;\eta] \;=[t\xi_1,\cdots,t\eta_L],$$
$$[\xi;\eta]^t=[\xi_1{}^t,\cdots,\eta_L{}^t].$$

We shall on occasion denote matrices of solutions by German capitals.

It is immediately evident that

the product of a solution of (S) *and a solution of* (M) *is a solution of* (S);
the product of two solutions of (M) *is a solution of* (M);
the sum of two solutions of (A) *is a solution of* (A).

The solution $x_1=x_2=\cdots=x_K=y_1=y_2=\cdots=y_L=1$ of (M) will be called the trivial solution of (M) and denoted by

$$\mathfrak{I}=[1;1].$$

The trivial solution of (A) is defined in an analogous manner as

$$\mathfrak{O}=[0;0].$$

A solution of (A) is said to be irreducible if it cannot be expressed as the sum of two non-trivial solutions *; similarly, a solution of (M) is said to be irreducible if it cannot be expressed as the product of two non-trivial solutions. Lastly, a solution of (S) is said to be irreducible if it cannot be expressed as the product of a solution of (S) and a non-trivial solution of (M).

The Greek letters α and β appearing as sub-scripts or super-scripts will have the ranges $1,2,\cdots,K$ and $1,2,\cdots,L$ respectively. Thus we write $x_\alpha=P^{u_\alpha}$ for $x_1=P^{u_1}$, $x_2=P^{u_2},\cdots,x_K=P^{u_K}$,

$$\sum_{(\beta)} v_\beta \text{ for } v_1+v_2+\cdots+v_L, \qquad \prod_{(\alpha)} P^{u_\alpha} \text{ for } P^{u_1}P^{u_2}\cdots P^{u_K},$$

and so on.

3. We shall first give some properties of the system (M).

THEOREM 3.1. *Every primitive solution of* (M) *is of the form*

$$x_\alpha=P^{u_\alpha}, \qquad y_\beta=P^{v_\beta}$$

where P is a prime, and $[u;v]$ is a primitive solution of (A).

* Grace and Young, p. 102.

Proof. Assume that (M) has a primitive solution $[\xi;\eta]$. Then there exists a prime P dividing at least one of the numbers ξ, η. Write

$$\xi_a = P^{u_a}\xi'^a, \qquad \eta_\beta = P^{v_\beta}\eta'_\beta$$

where the ξ', η' are prime to P. Substituting these numbers in (M), we obtain

$$\prod_{(a)} P^{a_{ia}u_a} \prod_{(a)} \xi'^{a_{ia}} = \prod_{(\beta)} P^{b_{i\beta}v_\beta} \prod_{(\beta)} \eta'^{b_{i\beta}}, \qquad (i = 1, \cdots, M).$$

Therefore

$$(3.1) \qquad \prod_{(a)} P^{a_{ia}u_a} = \prod_{(\beta)} P^{b_{i\beta}v_\beta}, \qquad \prod_{(a)} \xi'^{a_{ia}} = \prod_{(\beta)} \eta'^{b_{i\beta}}, \qquad (i = 1, \cdots, M),$$

and $[P^u; P^v]$, $[\xi'; \eta']$ are solutions of (M). Since the first is non-trivial, the second must be trivial, and

$$[\xi;\eta] = [P^u; P^v].$$

$[u; v]$ must be a primitive solution of (A). For from the first set of equations in (3.1)

$$\sum_{(a)} a_{ia}u_a = \sum_{(\beta)} b_{i\beta}v_\beta, \qquad (i = 1, \cdots, M),$$

so that $[u; v]$ is a solution of (A). But if it were the sum of two non-trivial solutions of (A), $[P^u; P^v]$ would be the product of two non-trivial solutions of (M).

COROLLARY. *Both the systems* (A) *and* (M) *have non-trivial solutions, or both have only trivial solutions.*

The primitive solution $[P^u; P^v]$ of (M) will be said to be of type $[u; v]$. There are an infinite number of primitive solutions of (M) of a given type; namely, one for each rational prime P. However *the number of types of primitive solutions of* (M) *is finite,* for the number of primitive solutions of (A) is known to be finite.*

Suppose that (A) has in all the R distinct primitive solutions

$$\mathfrak{U}_i = [\xi_i; \eta_i], \qquad (i = 1, 2, \cdots, R).$$

THEOREM 3.2. *Every solution of* (M) *is of the form*

$$(3.2) \qquad \begin{aligned} x_a &= T_1^{\xi_{1a}} T_2^{\xi_{2a}} \cdots T_R^{\xi_{Ra}} \\ y_\beta &= T_1^{\eta_{1\beta}} T_2^{\eta_{2\beta}} \cdots T_R^{\eta_{R\beta}} \end{aligned}$$

where the parameters $T_1, T_2, \cdots, T_R$ are positive integers. Conversely, every such expression is a solution of (M).

---

* Grace and Young, p. 103.

Proof. From the proof of theorem 3.1, it is evident that any solution $[\lambda;\mu]$ of (M) is of the form

$$\prod_{\sigma=1}^{S}[P_\sigma^{u_{\sigma1}}, P_\sigma^{u_{\sigma2}},\cdots, P_\sigma^{u_{\sigma K}}; P_\sigma^{v_{\sigma1}}, P_\sigma^{v_{\sigma2}},\cdots, P_\sigma^{v_{\sigma L}}],$$

where $P_1, P_2,\cdots, P_S$ are the distinct primes dividing $\lambda_1\lambda_2\cdots\lambda_K\,\mu_1\mu_2\cdots\mu_L$, and the $[u_\sigma; v_\sigma]$ are solutions of (A). Now *

$$[u_\sigma; v_\sigma] = k_1^{(\sigma)}\mathfrak{U}_1 + k_2^{(\sigma)}\mathfrak{U}_2 + \cdots + k_R^{(\sigma)}\mathfrak{U}_R$$

where the $k^{(\sigma)}$ are non-negative integers. Therefore

$$P_\sigma^{u_{\sigma\alpha}} = \prod_{\tau=1}^{R} P^{k_\tau^{(\sigma)}\xi_{\tau\alpha}}, \qquad P_\sigma^{v_{\sigma\beta}} = \prod_{\tau=1}^{R} P_\sigma^{k_\tau^{(\sigma)}\eta_{\tau\beta}}.$$

Accordingly,

$$\lambda_\alpha = \prod_{(\sigma)} P_\sigma^{u_{\sigma\alpha}} = \prod_{(\sigma)}\prod_{(\tau)} P_\sigma^{k_\tau^{(\sigma)}\xi_{\tau\alpha}} = \prod_{(\tau)}\prod_{(\sigma)} P_\sigma^{k_\tau^{(\sigma)}\xi_{\tau\alpha}} = \prod_{(\tau)} T_\tau^{\xi_{\tau\alpha}},$$

$$\mu_\beta = \prod_{(\sigma)} P_\sigma^{v_{\sigma\beta}} = \prod_{(\sigma)}\prod_{(\tau)} P_\sigma^{k_\tau^{(\sigma)}\eta_{\tau\beta}} = \prod_{(\tau)}\prod_{(\sigma)} P_\sigma^{k_\tau^{(\sigma)}\eta_{\tau\beta}} = \prod_{(\tau)} T_\tau^{\eta_{\tau\beta}},$$

where

$$T_\tau = \prod_{(\sigma)} P_\sigma^{k_\tau^{(\sigma)}} = P_1^{k_\tau^{(1)}} P_2^{k_\tau^{(2)}}\cdots P_S^{k_\tau^{(s)}}, \qquad (\tau = 1, 2,\cdots, R),$$

so that the T are positive integers. The converse of the theorem is obvious from the relations just given.

4. Since for each fixed value of α there must be at least one value of τ for which $\xi_{\tau\alpha} \neq 0$, and for each fixed value of β one value of τ for which $\eta_{\tau\beta} \neq 0$, none of the parameters T in (3.2) can be equal to unity for all solutions of (M) unless all solutions of (M) are trivial. In other words, *the number of primitive solutions of* (A) *gives the minimum number of parameters T necessary to express every solution of* (M) *in the form* (3.2).

The question naturally arises whether we can determine this number *a priori* without actually exhibiting all the primitive solutions of (A). In general, this appears to be impossible, but there are certain fairly general special systems (M) for which such a determination can be made. We give in this connection the following two theorems.

THEOREM 4.2. *The total number of parameters T necessary to express all solutions of the system*

$$(\mathrm{M}') \qquad x_1^{a_1}x_2^{a_2}\cdots x_K^{a_K} = y_{11}y_{12}\cdots y_{1L_1} = \cdots = y_{n1}y_{n2}\cdots y_{nL_n}$$

is given by the formula

$$\sum_{\alpha=1}^{K}\prod_{\tau=1}^{n}\binom{L_\tau + a_\alpha - 1}{a_\alpha}.$$

* Grace and Young, pp. 104, 103.

Here $\binom{m}{n}$ denotes as usual the number of combinations of m things taken n at a time.

THEOREM 4. 2. *The total number of parameters T necessary to express all solutions of the system*

$$(\mathrm{M}'') \quad (x_{11}x_{12}\cdots x_{1K_1})^{a_1} = (x_{21}x_{22}\cdots x_{2K_2})^{a_2} = \cdots = (x_{n1}x_{n2}\cdots x_{nK_n})^{a_n}$$

is given by the formula

$$\prod_{\tau=1}^{n}\binom{a'_\tau + K_\tau - 1}{a'_\tau},$$

where $a'_\tau = a/a_\tau$, $(\tau = 1, 2, \cdots, n)$, and a is the least common multiple of integers $a_1, a_2, \cdots, a_n$.

To illustrate these theorems,* consider the three systems

$$\text{(i)} \qquad x^2y^3z^3 = uv \;\;= wrst,$$

$$\text{(ii)} \qquad x^3y^3z^3 = u^2v^2 = wrst,$$

$$\text{(iii)} \qquad x^9 = y^5 = u^4v^4 = wrst.$$

For the first system, we apply theorem 4. 1 with $K = 3$, $a_1 = 2$, $a_2 = a_3 = 3$, $n = 2$, $L_1 = 2$, $L_2 = 4$,

$$\sum_{a=1}^{3}\prod_{\tau=1}^{2}\binom{L_\tau + a_a - 1}{a_a} = \sum_{a=1}\binom{a_a + 1}{a_a}\binom{a_a + 3}{a_a} = \binom{3}{2}\binom{2}{5} + 2\binom{4}{3}\binom{6}{3} = 190.$$

For the second system, we apply theorem 4. 2 wth $n = 3$, $a_1 = 3$, $a_2 = 2$, $a_3 = 1$, $K_1 = 3$, $K_2 = 2$, $K_3 = 4$, $a = 6$, $a'_1 = 2$, $a'_2 = 3$, $a'_3 = 6$,

$$\prod_{\tau=1}^{3}\binom{a'_\tau + K_\tau - 1}{a'_\tau} = \binom{4}{2}\binom{4}{3}\binom{9}{6} = 2{,}016.$$

For the third system, which involves only five algebraically independent variables, theorem 4. 2 gives

$$\prod_{\tau=1}^{4}\binom{K_\tau + a'_\tau - 1}{a'_\tau} = \binom{20}{20}\binom{36}{36}\binom{46}{45}\binom{183}{180} = \binom{46}{1}\binom{183}{3} = 46{,}217{,}626.$$

From these illustrations it is clear that even for rather simple looking

* In the last section of the paper will be found a simple system for which a verification of the theorems is feasible. If we take in (M') $a_1 = a_2 = \cdots = a_k = 1$ or in (M'') $a_1 = a_2 = \cdots = a_n = 1$, we find that the total number of parameters necessary to express all solutions of the system

$$x_{11}\,x_{12}\cdots x_{1k_1} = x_{21}\,x_{22}\cdots x_{2k_2} = \cdots = x_{n1}\,x_{n2}\cdots x_{nk_n}$$

is $\displaystyle\sum_{a=1}^{k_1}\prod_{\tau=a}^{n}\binom{K_\tau}{1} = \prod_{\tau=1}^{n}\binom{K_\tau}{1} = k_1\cdot k_2\cdots k_n$, a result obtained by Bell in the paper already cited by an entirely different argument.

systems, the number of parameters may be extraordinarily large, and that the actual exhibition of the solutions of a given system in the form (3.2) is usually impracticable.

The proof of theorem 4.1 is as follows. Consider the additive system associated with (M′),

$$(\mathrm{A}') \quad a_1 z_1 + \cdots + a_K z_K = w_{11} + \cdots + W_{1L_1} = \cdots = w_{n1} + \cdots + w_{nL_n}.$$

We have seen that the number of parameters T necessary for the solution of (M′) is the number of primitive solutions of (A′).

There exist solutions of (A′) with one of the z equal to one and all the remaining z equal to unity, and every such solution is primitive. Let us consider those solutions in which $z_a = 1$ and $z_1 = z_2 = \cdots = z_{a-1} = z_{a+1} = \cdots = z_K = 0$.

For such a solution we must have from (A′) n relations of the type

$$a_a = w_1 + w_2 + \cdots + w_L \tag{4.1}$$

where the w are non-negative integers. But the total number of ways that we can choose such numbers w to satisfy (4.1) equals the coefficient of t^{a_a} in the product $(1 + t + t^2 + \cdots)^L$, which is $\binom{L + a_a - 1}{a_a}$.

Therefore the total number of solutions under consideration is

$$\prod_{\tau=1}^{n} \binom{L_\tau + a_a - 1}{a_a} .$$

On taking $\alpha = 1, 2, \cdots, K$ it follows that *there are at least*

$$\sum_{a=1}^{k} \prod_{\tau=1}^{n} \binom{L_\tau + a_a - 1}{a_a} \text{ primitive solutions of } (\mathrm{A}').$$

To show that there are exactly this number, it suffices to show that no solution of (A′) not of the special form considered can be primitive.

Let the values of z in such a solution be $\eta_1, \eta_2, \cdots, \eta_K$ where $\eta_i \neq 0$ and let $N = a_1\eta_1 + a_2\eta_2 + \cdots + a_K\eta_K$, $M = a_i$. Then by our hypothesis, $N > M$.

It follows as for (4.1) that the values of w in any one of the sums in (A′) must form a partition of N into L or fewer parts. But for every such partition of N,

$$N = \gamma_1 + \gamma_2 + \cdots \gamma_{L'},$$

where $\gamma_1 \geqq \gamma_2 \geqq \cdots \geqq \gamma_{L'} > 0$, $(L' \leqq L)$ we can find a partition of M

$$M = \partial_1 + \partial_2 + \cdots + \partial_{K'}$$

such that $K' \leqq L'$, $\partial_j \leqq \gamma_j$, $(j = 1, 2, \cdots, K')$.

Therefore by assigning the proper w to the γ and ∂, we exhibit our solution as the sum of a primitive solution of (A′) and a non-trivial solution of (A′) associated with a certain set of partitions of $N-M$.

The proof of theorem 4.2 follows similar lines. With an obvious extension of our matrix notation, let

$$[\xi^{(1)};\xi^{(2)};\cdots;\xi^{(n)}]$$

be a solution of the additive system associated with (M″),

$$(\text{A}'') \qquad a_1(z_{11}+\cdots+z_{1K_1})=\cdots=a_n(z_{n1}+\cdots+z_{nK_n})$$

and let

$$N_\tau=\xi_1^{(\tau)}+\xi_2^{(\tau)}+\cdots+\xi_{K_\tau}^{(\tau)}, \qquad (\tau=1,2,\cdots,n).$$

Then

$$(4.2) \qquad a_1N_1=a_2N_2=\cdots=a_nN_n=N, \quad \text{say.}$$

Now for integral $N_1,\cdots,N_n$ the least positive value of N which can satisfy a relation of the form (4.2) is the least common multiple of $a_1,a_2,\cdots,a_n$. Denote this number by a, and let

$$a'_\tau=a/a_\tau, \qquad (\tau=1,2,\cdots,n).$$

Then if

$$(4.3) \qquad a'_\tau=\eta_1^{(\tau)}+\eta_2^{(\tau)}+\cdots+\eta_{K_\tau}^{(\tau)}$$

is a partition of a'_τ into K_τ parts, zero counting as a part,

$$[\eta^{(1)};\eta^{(2)};\cdots;n^{(n)}]$$

will be a primitive solution of (A″). There are $\binom{a'_\tau+K_\tau-1}{a'_\tau}$ distinct ways of selecting non-negative $\eta^{(\tau)}$ to satisfy (4.3), and hence in all

$$\prod^{n}\binom{a'_\tau+K_\tau-1}{a'_\tau}$$

such primitive solutions. The proof that there are no other primitive solutions is almost exactly the same as for Theorem 4.1.

5. The results of section three allow us to complete the discussion of the general system (S).

Let $P_1,P_2,\cdots,P_H$ be the distinct prime factors of the $2M$ integers $A_1,\cdots,B_M$ so that

$$A_i=P_1^{c_{i1}}P_2^{c_{i2}}\cdots P_H^{c_{iH}}, \quad B_i=P_1^{d_{i1}}P_2^{d_{i2}}\cdots P_H^{d_{iH}}, \quad (i=1,\cdots,M)$$

where the c and d are non-negative integers, and for a fixed k at least one of the $2M$ numbers $c_{1k},c_{2k},\cdots,c_{Mk},d_{1k},d_{2k},\cdots,d_{Mk}$ is positive.

Consider the system

$$(M^{(k)}) \quad P_k^{c_{ik}} x_1^{a_{i1}} \cdots x_K^{a_{iK}} = P_k^{d_{ik}} y_1^{b_1} \cdots y_L^{b_{iL}}, \qquad (i = 1, \cdots, M)$$

and the associated additive system

$$(A^{(k)}) \quad c_{ik} + a_{i1} z_1 + \cdots + a_{iK} z_K = d_{ik} + b_{i1} w_1 + \cdots + b_{iL} w_L, \qquad (i = 1, \cdots, M).$$

Then if a primitive solution of $(A^{(k)})$ is defined as one which cannot be expressed as the sum of a solution of $(A^{(k)})$ and a non-trivial solution of (A), it follows as in the proof of theorem 3.1 that every primitive solution of $(M^{(k)})$ is of the form $[P_k^{\lambda}; P_k^{\mu}]$ where $[\lambda; \mu]$ is a primitive solution of $(A^{(k)})$.

Consider in connection with $(A^{(k)})$ the additive system

$$(B^{(k)}) \quad c_{ik} z_0 + a_{i1} z_1 + \cdots + a_{iK} z_K = d_{ik} w_0 + b_{i1} w_1 + \cdots + b_{iL} w_L, \qquad (i = 1, \cdots, M).$$

Then the number of primitive solutions of $(B^{(k)})$ is finite. If among these primitive solutions there are l_0 with $z_0 = w_0 = 1$, l_1 with $z_0 = 0$, $w_0 = 1$ and l_2 with $z_0 = 1$, $w_0 = 0$ then $(A^{(k)})$ and hence $(M^{(k)})$ has exactly $\nu_k = l_0 + l_1 l_2$ primitive solutions. If $l_0 + l_1 l_2 = 0$, $(M^{(k)})$ has no primitive solutions, and hence no solutions whatever. We shall see in a moment that this would entail (S) having no solutions contrary to our hypothesis. Hence $\nu_k > 0$ and the primitive solutions of $(M^{(k)})$ may be exhibited, since the primitive solutions of $(B^{(k)})$ can be found by trial in a finite number of steps.*

If we denote such a primitive solution of $(M^{(k)})$ by $[\xi^{(k)}; \eta^{(k)}]$, then

$$(5.1) \qquad [\xi; \eta] = [\xi^{(1)}; \eta^{(1)}] \cdot [\xi^{(2)}; \eta^{(2)}] \cdots [\xi^{(H)}; \eta^{(H)}]$$

is a primitive solution of (S), and there are in all exactly $\nu = \nu_1 \nu_2 \cdots \nu_H$ such solutions. Conversely, if (S) has solutions, and hence primitive solutions, a decomposition such as (5.1) is possible, so that each $(M^{(k)})$ must have primitive solutions. We summarize our results in the following theorem.

THEOREM 5.1. *If* (S) *has solutions, every solution is of the form*

$$(5.2) \qquad \begin{aligned} x_a &= C_a T_1^{\xi_{1a}} T_2^{a\xi_2} \cdots T_s^{\xi_{sa}} \\ y_\beta &= D_\beta T_1^{\eta_{1\beta}} T_2^{\eta_{2\beta}} \cdots T_s^{\eta_{s\beta}} \end{aligned}$$

where the T, ξ and η are as in Theorem 3.2, and the pairs of integers C_a, D_β may assume at most ν sets of values, where ν is given in the discussion above.

6. We have not treated here the important problem of what restrictions

---

* Grace and Young, p. 104.

it is necessary to impose upon the parameters T so that the formulas (5.1) shall give the solutions of (S) once and once only.* This question is bound up in a highly interesting manner with the co-primality of sets of the parameters and their restriction to be numbers of a special form; e. g. square free. I hope to give some results connected with this problem subsequently.

I conclude by solving by the additive method the system used by Bell to illustrate his general process of solution,†

(iv) $$x_1x_2^2 = y_1y_2 = z_1z_2.$$

The additive dual of (iv) is

(6.1) $$X_1 + 2X_2 = Y_1 + Y_2 = Z_1 + Z_2.$$

By inspection we can write down the following thirteen primitive solutions of (6.1):

$$\begin{aligned}
\mathfrak{U}_1 &= [1,0;\ 1,0;\ 1,0], & \mathfrak{U}_7 &= [0,1;\ 0,2;\ 2,0],\\
\mathfrak{U}_2 &= [1,0;\ 1,0;\ 0,1], & \mathfrak{U}_8 &= [0,1;\ 0,2;\ 0,2],\\
\mathfrak{U}_3 &= [1,0;\ 0,1;\ 1,0], & \mathfrak{U}_9 &= [0,1;\ 1,1;\ 2,0],\\
\mathfrak{U}_4 &= [1,0;\ 0,1;\ 0,1], & \mathfrak{U}_{10} &= [0,1;\ 1,1;\ 0,2],\\
\mathfrak{U}_5 &= [0,1;\ 2,0;\ 2,0], & \mathfrak{U}_{11} &= [0,1;\ 2,0;\ 1,1],\\
\mathfrak{U}_6 &= [0,1;\ 2,0;\ 0,2], & \mathfrak{U}_{12} &= [0,1;\ 0,2;\ 1,1],
\end{aligned}$$
$$\mathfrak{U}_{13} = [0,1;\ 1,1;\ 1,1].$$

By theorem (4.1), the solution of (iv) will contain $\binom{2}{1}\binom{2}{1} + \binom{3}{2}\binom{3}{2} = 13$ parameters.

Hence $\mathfrak{U}_1, \cdots, \mathfrak{U}_{13}$ are all the primitive solutions of (6.1), so that by theorem (3.2) the solution of (iv) is

$$\begin{aligned}
x_1 &= T_1T_2T_3T_4, & x_2 &= T_5T_6T_7T_8T_9T_{10}T_{11}T_{12}T_{13},\\
y_1 &= T_1T_2T_5^2T_6^2T_9T_{10}T_{11}^2T_{13}, & y_2 &= T_3T_4T_7^2T_8^2T_9T_{10}T_{12}^2T_{13},\\
z_1 &= T_1T_3T_5^2T_7^2T_9^2T_{11}T_{12}T_{13}, & z_2 &= T_2T_4T_6^2T_8^2T_{10}^2T_{11}T_{12}T_{13}.
\end{aligned}$$

On making the change of variables

$$\begin{matrix} T_1, & T_2, & T_3, & T_4, & T_5, & T_6, & T_7, & T_8, & T_9, & T_{10}, & T_{11}, & T_{12}, & T_{13} & \text{into} \\ \phi_1, & \phi_2, & \phi_3, & \phi_4, & \psi_9, & \psi_5, & \psi_4, & \psi_8, & \psi_7, & \psi_6, & \psi_2, & \psi_1, & \psi_3, & \end{matrix}$$

this solution agrees with that obtained by Bell.

The additive method gives no information about the co-primeness of the parameters T, and it is to some extent tentative. In compensation, it is usually shorter that the multiplicative method.

---

* See Elliott, *Quarterly Journal of Mathematics*, Vol. 34 (1903), pp. 348-377 for a discussion of the similar problem for (A) in the case $M = 1$, with considerable detail for the sub-case $K + L = 3$.

† Paper cited, § 15.

CONCERNING PRIMITIVE GROUPS OF CLASS *U*.

By C. F. Luther.

1. It is the purpose of this paper to develop limits to the degree n of multiply transitive groups of class u (> 3) that contain substitutions of degree $u + \epsilon$, ϵ a positive integer, these substitutions having certain restrictions upon their order. Use is made of the effective methods devised by Bochert * and Manning † in their work upon the class problem. Three theorems are to be proved.

Theorem I. *If n is the degree and u is the class of a group that contains a substitution of order 2 and degree $u + \epsilon$, ϵ a positive integer, then if the group is*

doubly transitive,	$u > n/2 - n^{1/2}/2 - 5\epsilon$,	*provided* $\epsilon < u/5$
triply transitive,	$n \leqq 2u + 8\epsilon$,	*provided* $\epsilon < u/6$
5-ply transitive,	$n < 5u/3 + 5\epsilon$,	*provided* $\epsilon < u/6$
6-ply transitive,	$n < 4u/3 + 4\epsilon$,	*provided* $\epsilon < u/9$
7-ply transitive,	$n < 13u/10 + 4\epsilon$,	*provided* $\epsilon < u/8$
8-ply transitive,	$n < 6u/5 + 4\epsilon$,	*provided* $\epsilon < u/8$
9-ply transitive,	$n < 8u/7 + 6\epsilon$,	*provided* $\epsilon < u/10$
11-ply transitive,	$n < 12u/11 + 5\epsilon$,	*provided* $\epsilon < u/10$

more than p (a prime $\geqq 11$)-ply transitive,

$$n < (p-1)u/(p-2) + 3\epsilon, \quad \textit{provided} \quad u > p + 7\epsilon/2$$

more than σ $(= p_1 + p_2 + p_3 + \cdots + p_r \geqq 13)$ times transitive, where $p_1, p_2, p_3, \cdots, p_r$ are distinct odd primes and $r > 1$,

$$u > n - \epsilon - \frac{n + 2\epsilon \prod_1^r (p_k)}{\prod_1^r (p_k - 1)}, \qquad \textit{provided}$$

$$u > \sigma + \frac{4 \prod_1^r (p_k)^2}{\prod_1^r (p_k - 1)^2} \left(\epsilon + \frac{\sigma}{\prod_1^r (p_k)} \right) \quad \text{and} \quad r \leqq \frac{4 \prod_1^r (p_k)}{\prod_1^r (p_k - 1)} \epsilon.$$

$$\textit{If } r \geqq \frac{2 \prod_1^r (p_k)}{\prod_1^r (p_k - 1)} \epsilon, \quad u > n - 2\epsilon - r - \frac{n}{\prod_1^r (p_k - 1)}.$$

---

* Bochert, *Mathematische Annalen*, Vol. 40 (1892), pp. 176-193; and Vol. 49 (1897), pp. 133-144.

† Manning, *Transactions of the American Mathematical Society*, Vol. 18 (1917), pp. 463-479; and Vol. 31 (1928), pp. 643-653.

2. It is necessary first to prove an auxiliary theorem concerning diedral rotation groups:

If the order of a group of class equal to or greater than u, generated by two substitutions s and t of order 2 and degree $u+\epsilon$, ϵ a positive integer, is divisible by each of the distinct odd primes $p_1, p_2, \cdots, p_r$, its degree n is less than

$$u+(u+2\epsilon p_1 p_2 p_3 \cdots p_r)/(p_1-1)(p_2-1)\cdots(p_r-1).$$

A corresponding theorem has been proved by Professor Manning * for the special case $\epsilon=0$. The method of proof here is in general similar to his and so will be given only briefly.

Consider a group generated by s and t. Professor Manning has shown that if it contains regular constituents involving ξ letters displaced by both s and t, we can erase these constituents and consider the resulting group of degree $n-\xi$, generated by two substitutions of order 2 and degree $u+\epsilon-\xi$. If the theorem is true under such conditions, it is likewise true when the regular constituents are present.

Assume s has m_1 cycles that displace letters not in t, and that t has m_2 cycles that displace letters not in s. There are y_i constituents of degree Y_i and order $2Y_i$, Y_i an odd number and $(i=1,2,3,\cdots)$. There are z'_k constituents of degree Z_k and order $2Z_k$, Z_k an even number, with the generator of degree Z_k in s and that of degree Z_k-2 in t. Similarly there are z_k'' constituents of the same degree Z_k, with Z_k-2 letters in s and Z_k letters in t. Hence the degree of s is

$$(1)\qquad 2m_1+\Sigma y_i(Y_i-1)+\Sigma z'_k Z_k+\Sigma z_k''(Z_k-2),$$

and the degree of t is

$$(2)\qquad 2m_2+\Sigma y_i(Y_i-1)+\Sigma z'_k(Z_k-2)+\Sigma z_k'' Z_k, \qquad (i,k=1,2,\cdots).$$

Let $$E_a\equiv\Sigma y_i(Y_i-a);\quad F_a\equiv\Sigma z_k(Z_k-a);\quad \Gamma_a\equiv E_a+F_a;$$
$$m\equiv m_1+m_2;\quad z\equiv z'_k+z_k''.$$

From (1) and (2) we have

$$(3)\qquad m+\Gamma_1=u+\epsilon,$$

and since the degree n of $\{s,t\}$ is $2m+\Sigma y_i Y_i+\Sigma z_k Z_k$,

$$(4)\qquad n=2m+\Gamma_0,$$

and from (3) and (4)

* Manning, *Transactions of the American Mathematical Society*, Vol. 18 (1917), pp. 464 ff.

(5) $$n = 2(u+\epsilon) - \Gamma_2.$$

3. Now consider the order N of the product st:

Case 1. If N is of odd prime order,

$$n = u + (u + \epsilon p)/(p-1).$$

If $N = p^a$, p any prime and $\alpha > 1$, we consider $(st)^{p^k}$, $(k = 0, 1, 2, \cdots, \alpha - 1)$. If the group $\{s, t\}$ has x_i transitive constituents of degree p^i, $(i = 1, 2, \cdots, \alpha)$, we have α equations:

$$\begin{aligned} p^a x_a &= u + h_a, \\ p^{a-1}x_{a-1} + p^a x_a &= u + h_{a-1}, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ &\cdot\ \cdot\ \cdot\ \cdot\ , \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ &\cdot\ \cdot\ \cdot\ \cdot\ , \\ p^2x_2 + p^3x_3 + \cdots + p^a x_a &= u + h_2, \\ px_1 + p^2x_2 + p^3x_3 + \cdots + p^a x_a &= u + \delta. \end{aligned}$$

In addition we use formula (3) to give us $\alpha + 1$ consistent equations, from which we find that

$$n \leqq u + (u + \epsilon p^a)/p^{a-1}(p-1).$$

Case 2. If the order of st is twice the power of an odd prime, we find by a similar method that

$$n \leqq u + (u + 2p^a\epsilon - \epsilon)/(p^a - 1).$$

Case 3. If the order of st is the product of r distinct odd prime factors $p_1, p_2, p_3, \cdots, p_r$, then $m = 0$, $F_1 = 0$, and (3) becomes $E_1 = u + \epsilon$. In addition to this last equation there are $2^r - 1$ equations set up from the $N/p_ip_j \cdots$ powers of st. These 2^r equations in the $2^r - 1$ variables are consistent, and the determinant of the set must vanish. We have

$$n \leqq u + (u + \epsilon p_1 \cdot p_2 \cdots p_r)/(p_1 - 1)(p_2 - 1) \cdots (p_r - 1).$$

Case 4. If the order of st is twice the product of two distinct odd primes p and q, we can show that

$$n \leqq u + (u + 2pq\epsilon)/q(p-1).$$

Case 5. Let the order of st contain two prime power factors p^a and q^β, p and q distinct odd primes. Then $N = p^a q^\beta \phi$, where ϕ is unity or a number relatively prime to p and q. In this case

$$n \leqq u + 2u/p^a q^\beta + 2\epsilon.$$

Case 6. If the order N of st is $\pi\phi$, where π $(= p_1^{a_1} \cdot p_2^{a_2} \cdots p_r^{a_r})$ is

the product of powers of r (>2) distinct odd prime factors and ϕ is unity or relatively prime to π,

$$n \leqq u + 2u/\pi + 2\epsilon.$$

Case 7. If st is of order 2π, $\pi = 3p^a$, p^a a prime power factor,

$$n \leqq u + u/\pi + 2u/\pi^2 + 2\epsilon.$$

Case 8. If st is of order 2π, where $\pi = p_1^{a_1} \cdot p_2^{a_2} \cdots p_r^{a_r}$, each factor greater than 3,

$$n \leqq u + u/\pi + 2u/\pi^2 + 2\epsilon.$$

We wish to replace these special limits for n, each depending upon the order of st, by the single limit

$$n < u + (u + 2\epsilon p_1 \cdot p_2 \cdots p_r)/(p_1 - 1)(p_2 - 1) \cdots (p_r - 1),$$

which covers all cases.

4. We proceed now to the proof of Theorem I. Begin by considering a p times transitive group, p an odd prime.* By hypothesis there is a substitution in the group G of order 2 and degree $u + \epsilon$. It is:

$$S = (a_0a_2)(a_3a_4) \cdots (a_{p-2}a_{p-1}) \cdots (a_p) \cdots.$$

Since G is p times transitive, there must be

$$S' = (a_0a_3)(a_2a_5)(a_4a_7) \cdots (a_{p-3}\cdot) \cdots (a_{p-1}),$$

similar to S. Transform S' by the g_{p-1} substitutions of the transitive subgroup that fixes $a_0, a_2, \cdots, a_{p-1}$, and obtain a set of g_{p-1} similar substitutions $S', S'', \cdots, S^{(g_{p-1})}$. $S^{(i)}$ has in common with S the $p-2$ letters $a_0, a_2, \cdots, a_{p-2}$, and x_i (unknown) other letters (in common). We shall calculate Σx_i for the set of g_{p-1} similar substitutions. There are $(u + \epsilon - p + 1)g_{p-1}/(n - p + 1)$ substitutions that contain a given letter not $a_0, \cdots, a_{p-1}$. There are but $u + \epsilon - p + 1$ letters that can be called common letters. Hence

$$\Sigma x_i = (u + \epsilon - p + 1)\,\frac{(u + \epsilon - p + 2)}{n - p + 1}\,g_{p-1}.$$

There are exactly $(n - u - \epsilon)g_{p-1}/(n - p + 1)$ substitutions that replace a_{p-3} by a letter new to S, and $(u + \epsilon - p + 1)g_{p-1}/(n - p + 1)$ substitutions that replace a_{p-3} by a letter of S. In the first case the product of any one of these substitutions $S^{(i)}$ with S contains a cycle of order p. Our auxiliary theorem says that

---

* Cf. Manning, *Transactions of the American Mathematical Society*, Vol. 18 (1917), pp. 474 ff.

$$x_i > [u(p-2) - 2\epsilon - (p-2)(p-1)]/(p-1).$$

Hence in all the $(n-u-\epsilon)g_{p-1}/(n-p+1)$ substitutions

$$\Sigma x_i > \frac{u(p-2) - 2\epsilon - (p-2)(p-1)}{p-1} \cdot \frac{(n-u-\epsilon)}{n-p+1} g_{p-1}.$$

For the substitutions $S^{(4)}$ that replace a_{p-3} by a letter of S, the product $SS^{(4)}$ will contain a cycle of order $p+1$ or greater. We are interested in the primes 3, 5, and 7. The most unfavorable case from the standpoint of common letters is when $N = 2 \cdot 3$, in which case the least number of letters of $S^{(4)}$ in common with S is

$$u/2 - \epsilon/2 - 1.$$

Hence

$$\Sigma x_i \geqq \left(\frac{u-\epsilon}{2} - 1\right) \frac{(u+\epsilon-p+1)}{n-p+1} g_{p-1},$$

where the right-hand member represents the minimum number of common letters in these substitutions. Therefore, over the entire g_{p-1} substitutions

$$\Sigma x_i > \frac{n-u-\epsilon}{n-p+1} \cdot g_{p-1} \cdot \frac{(p-2)u - 2\epsilon - (p-2)(p-1)}{p-1}$$
$$+ \frac{u-\epsilon-2}{2} \cdot \frac{u+\epsilon-p+1}{n-p+1} g_{p-1},$$

and finally

$$\begin{aligned}(u+\epsilon-p+1)&(u+\epsilon-p+2)\\ &> (n-u-\epsilon)\cdot[(p-2)u - 2\epsilon - (p-2)(p-1)]/(p-1)\\ &\quad + [(u-\epsilon-2)/2](u+\epsilon-p+1),\end{aligned}$$

which becomes

$$(n-u-\epsilon) < \tfrac{1}{2}\, \frac{(u+\epsilon-p+1)(u+3\epsilon-2p+6)}{u - \dfrac{2\epsilon}{p-2} - p + 1} \cdot \frac{p-1}{p-2}.$$

When $p = 5$, $\quad n < 5u/3 + 5\epsilon$, provided $u > 6\epsilon$.

When $p = 7$, if a_4 is replaced by a letter of S, the order of $SS^{(4)}$ is not less than 8, excluding the possibility of $N = 6$. Hence

$$n < 13u/10 + 4\epsilon, \quad \text{provided} \quad u > 8\epsilon.$$

If $p = 3$, we have

$$(u+\epsilon-p+1)(u+\epsilon-p+2) \geqq (u/2 - \epsilon/2 - 1)(n-p+1)$$

from which

$$n \leqq 2u + 8\epsilon, \quad \text{provided} \quad u > 5\epsilon.$$

6

5. Now assume that G is more than p times transitive, p an odd prime. Then

$$S = (a_0a_2)(a_3a_4)\cdots(a_{p-2}a_{p-1})\cdots(a_p)\cdots$$

and

$$S' = (a_0a_3)(a_2a_5)\cdots(a_{p-3}a_p)\cdots(a_{p-1})\cdots.$$

Transform S' by the transitive subgroup fixing $a_0, a_2, \cdots, a_p$ to give a set of similar substitutions $S', S'', \cdots, S^{(g_p)}$, such that every product $SS^{(i)}$ is of order p or a multiple of p. Hence

$$\Sigma x_i > \frac{u(p-2) - 2\epsilon - (p-2)(p-1)}{p-1}\, g_p.$$

Also, any one of the last $u + \epsilon - p + 1$ letters of S is found in $(u + \epsilon - p + 1)g_p/(n-p)$ of the substitutions $S', S'', \cdots, S^{(g_p)}$. Hence

$$\Sigma x_i = \frac{(u + \epsilon - p + 1)^2}{n-p}\, g_p,$$

so

$$n - p < \frac{(u + \epsilon - p + 1)^2}{u - \dfrac{2\epsilon}{p-2} - p + 1}\ \frac{p-1}{p-2}.$$

Let $p = 5$ for 6-ply transitive groups, and we have

$$n < 4u/3 + 4\epsilon, \qquad \text{provided } u > 9\epsilon.$$

If $p = 7$,

$$n < 6u/5 + 4\epsilon, \qquad \text{provided } u > 8\epsilon.$$

In the general case of any prime p ($\geqq 11$)

$$n < (p-1)u/(p-2) + 3\epsilon, \qquad \text{provided } u > p + 7\epsilon/2,$$

or if $p \geqq 19$,

$$n > (p-1)u/(p-2) + 5\epsilon/2, \qquad \text{provided } u > p + 6\epsilon.$$

6. Let us now assume that G is more than σ times transitive, where σ is the sum of $r(>1)$ distinct odd primes $p_1, p_2, p_3, \cdots, p_r$. As before we have

$$S = (a_0a_2)\cdots(a_{p_1-2}a_{p_1-1})\cdots(c_0c_2)\cdots(c_{p_r-2}c_{p_r-1})\cdots(a_{p_1})\cdots(c_{p_r})\cdots$$

and

$$S' = (a_0a_3)\cdots(a_{p_1-3}a_{p_1})\cdots(c_0c_3)\cdots(c_{p_r-3}c_{p_r})\cdots(a_{p_1-1})\cdots(c_{p_r-1})\cdots.$$

Transform S' by the transitive subgroup G_σ that fixes the σ letters

$$a_0, a_2, \cdots, a_{p_1}, \cdots, c_0, c_2, \cdots, c_{p_r}.$$

Every product $SS^{(i)}$ is of order $\prod\limits_1^r (p_k)$ or a multiple of it. In the g_σ substitutions, $[u+\epsilon-\sum\limits_1^r (p_k-1)]\, g_\sigma/(n-\sigma)$ substitutions contain a given letter of G_σ. Hence

$$\Sigma x_i = \frac{[u+\epsilon-\Sigma(p_k-1)]^2}{n-\sigma}\, g_\sigma.$$

Our auxiliary theorem tells us that

$$x_i > \frac{\{\prod\limits_1^r (p_k-1)-1\}}{\prod\limits_1^r (p_k-1)}\, u + 2(r+\epsilon) - \frac{2\epsilon \prod\limits_1^r (p_k)}{\prod\limits_1^r (p_k-1)} - \sigma.$$

Hence over the g_σ conjugate substitutions,

$$\Sigma x_i > \left[\frac{\{\prod\limits_1^r (p_k-1)-1\}}{\prod\limits_1^r (p_k-1)}\, u + 2(r+\epsilon) - \frac{2\epsilon \prod\limits_1^r (p_k)}{\prod\limits_1^r (p_k-1)} - \sigma\right] g_\sigma.$$

Finally,

$$\text{(A)} \qquad n-\sigma < \frac{(u+\epsilon+r-\sigma)^2}{\dfrac{\phi}{\phi+1}\, u + 2r + 2\epsilon - 2\epsilon\lambda - \sigma},$$

where

$$\phi = \prod_1^r (p_k-1) - 1; \qquad \lambda = \prod_1^r (p_k) / \prod_1^r (p_k-1).$$

To determine a superior limit for n, we assume a relation

$$n < [(\phi+1)/\phi]u + k\epsilon.$$

In order that this may be a true relation, $n = (\phi+1)u/\phi + k\epsilon$ must fail to satisfy the preceding inequality. As k is at our disposal, we choose it with that in mind. After substituting in (A) and collecting terms we must have

$$\left[\left(\frac{k\phi}{\phi+1} + \frac{2}{\phi} - 2\lambda\frac{\phi+1}{\phi}\right)\epsilon + \frac{2r}{\phi} - \frac{\sigma}{\phi(\phi+1)}\right] u$$
$$+ [2k(1-\lambda)-1]\epsilon^2 + 2(k-1)r\epsilon + (2\lambda-k)\epsilon\sigma - r^2 > 0.$$

Let $k = 2\lambda + 1$, and replace $2r/\phi - \sigma/\phi(\phi+1)$ by the smaller quantity $(2\lambda+1)\sigma/\phi(\phi+1)$, which is true for primes excluding the sets 3, 5 and 3, 7. We now must have

$$\{[(\phi+1)^2 + 1 - 2\lambda(2\phi+1)]\,\epsilon + (2\lambda+1)\sigma\}u$$
$$> \phi(\phi+1)\{(4\lambda^2-2\lambda-1)\epsilon^2 - 4\lambda r\epsilon + \epsilon\sigma + r^2\}.$$

We wish now to determine the minimum value of u that will satisfy this inequality. Let $u = 4\lambda^2\epsilon + m\sigma$, and we must have

$$4\lambda^2 + 8\lambda^3 + [(\phi + 1)^2 + 1 - 2\lambda(2\phi + 1)]\, m > \phi(\phi + 1)$$

and

$$m\sigma^2 + (4\lambda r\epsilon - r^2)\phi(\phi + 1) > 0.$$

These are satisfied provided $m \geqq 1 + 4\lambda/(\phi + 1)$ and $r \leqq 4\lambda\epsilon$. Hence

$$n < [(\phi + 1)/\phi]u + (2\lambda + 1)\epsilon,$$

provided $u \geqq 4\lambda^2\epsilon + [1 + 4\lambda/(\phi + 1)]\sigma$ and $r \leqq 4\lambda\epsilon$. If $r \geqq 2\lambda\epsilon$, it is easily shown that we can always use

$$n < [(\phi + 1)/\phi]u + 2\epsilon + r.$$

In case of the primes 3, 5 and 3, 7, we find by direct substitution that

$$n < 8u/7 + 6\epsilon, \qquad \text{provided} \quad u > 10\epsilon$$

and

$$n < 12u/11 + 5\epsilon, \quad \text{provided} \quad u > 10\epsilon$$

respectively.

7. Consider now a doubly transitive group that contains a subs of order 2 and degree $u + \epsilon (u > 3)$. S is one of a set of w conjugates. e w conjugate substitutions contain $w(u + \epsilon)/2$ transpositions. A particular transposition occurs $w(u + \epsilon)/n(n - 1)$ times in the set. There are $(n - u - \epsilon)(u + \epsilon)$ possible combinations of one letter of S and one letter not of S. A substitution $S^{(i)}$ conjugate to S containing such a transposition is non-commutative with S. If y represents the number of such non-commutative conjugate substitutions, then

$$y \geqq 2w(u + \epsilon)(n - u - \epsilon)/n(n - 1).$$

The $(u + \epsilon)$ letters of S can be combined in pairs in $(u + \epsilon) \times (u + \epsilon - 1)/2$ different ways. The entire set of w substitutions has $w(u + \epsilon)(u + \epsilon - 1)/2$ of these pairs, and any one of the above pairs of letters of S occurs in $w(u + \epsilon)(u + \epsilon - 1)/n(n - 1)$ substitutions of the set. Summing over the w conjugates,

$$\sum_w x_i(x_i - 1) = w(u + \epsilon)^2(u + \epsilon - 1)^2/n(n - 1).$$

Also

$$\sum_w x_i = w(u + \epsilon)^2/n.$$

We make use of the identity,

$$\sum_w \left(x_i - \frac{1}{w} \sum_w x_i\right)^2 = \sum_w x_i^2 - \frac{1}{w}\left(\sum_w x_i\right)^2.$$

Hence

$$\sum_{w} [x_i - (u+\epsilon)^2/n]^2 = w(u+\epsilon)^2(u+\epsilon-1)^2/n(n-1) + w(u+\epsilon)^2/n - w(u+\epsilon)^4/n^2.$$

From the auxiliary formula, $x_i \geqq (u-\epsilon)/2$; then it follows that

$$\sum_{y} [x_i - (u+\epsilon)^2/n]^2 \geqq y[(u-\epsilon)/2 - (u+\epsilon)^2/n]^2.$$

We agree that $n > 2u + 8\epsilon$, $u > 3\epsilon$. Then

$$\sum_{w-y} [x_i - (u+\epsilon)^2/n]^2 \leqq w(u+\epsilon)^2(n-u-\epsilon)^2/n^2(n-1) - y[(u-\epsilon)/2 - (u+\epsilon)^2/n]^2;$$

but

$$\sum_{w-y} [x_i - (u+\epsilon)^2/n]^2 \geqq [1/(w-y)]\{\sum_{w-y} [x_i - (u+\epsilon)^2/n]\}^2;$$

hence

$$\sum_{w-y} [x_i - (u+\epsilon)^2/n]^2 \geqq [y^2/(w-y)]\,[(u-\epsilon)/2 - (u+\epsilon)^2/n]^2.$$

Finally,

$$\text{(B)}\quad \frac{(u+\epsilon)}{(u-\epsilon)^2}(n-u-\epsilon)\left(\frac{n}{2} - \frac{(u+\epsilon)(n-u-\epsilon)}{n-1}\right) \geqq \left(\frac{n}{2} - \frac{(u+\epsilon)^2}{u-\epsilon}\right)^2.$$

Replace $n/2 - (u+\epsilon)^2/(u-\epsilon)$ by $n/2 - u - 4\epsilon$ with the restriction that $u > 5\epsilon$, and assume a relation of the form

$$\text{(C)}\qquad u > n/2 - \gamma n^{1/2} - 4\epsilon - \delta$$

with γ and δ at our disposal. Then (B) becomes a polynomial in $n^{1/2}$ and for (C) to be a true relation, the inequality of (B) must fail to be satisfied. We pick $\gamma = 1/2$, to eliminate the highest power of n, and $\delta = \epsilon$. Direct calcualtion will show that (C) is true provided $n > 227$. Comparing (C) with Bochert's limit of $u > n/3 - 2n^{1/2}/3$ lowers the restriction to $n > 178$. Direct substitution into (B) verifies the truth of (C) except for the cases $\epsilon = 1, 2, 3$, when n must exceed 63, 69, and 87, respectively. A more careful study removes these exceptions also.

To demonstrate the method, take $\epsilon = 1$, $n > 63$. If $n = 63, 58, 56, 46$, (C) requires that $u > 22.5, 20.2, 19.3$, and 14.6, respectively. As ϵ is odd, u is odd. All doubly transitive groups of class 15 and less are known, and their degrees come well within the limit (C). We need only consider $u = 17$, 19, 21, 23. If $u = 17, 19, 23$, Jordan's theorem * restricts n to at most 19, 21, and 25, respectively. If $u = 21$, there is a substitution of degree 21 whose order is 7 or 3. The first case is covered by a theorem of Professor Manning †

* Jordan, *Traité des Substitutions*, 1870, p. 664.

† Manning, *Transactions of the American Mathematical Society*, Vol. 15 (1909), No. 2, p. 247.

which limits the degree to 24. In the second case it is possible to prove * that *if a doubly transitive group of class* 21 *contains a substitution of degree* 21 *and of order* 3, *its degree cannot exceed* 57.

8. It is well to recall that a simply transitive primitive group has no transitive subgroup of a lower degree. A doubly transitive group which is not triply transitive can have no doubly transitive subgroup of lower degree. We first prove that G can have a transitive subgroup H of degree at most 56. If H is of degree 56 and primitive, G is of degree at most 57. However, if H is imprimitive, further study is necessary to show that G cannot exceed 57. As this will be more easily understood when the structure of H is known, we leave it until later, and proceed to show that the degree of H cannot exceed 56.

First we assume that every substitution S_i of order 3 and degree 21 in G, that unites cycles of another substitution S_j of order 3 and degree 21 in G, has at least one cycle in which all the letters are new to S_j. Let

$$S_1 = (a_1a_2a_3)(b_1b_2b_3)(c_1c_2c_3)(d_1d_2d_3)(e_1e_2e_3)(f_1f_2f_3)(g_1g_2g_3),$$

and let S_2 be a substitution of order 3 and degree 21, that unites cycles of S_1 and has, by hypothesis, a cycle new to S_1. The transform $S_1^2S_2S_1$ has not two cycles new to S_2, for reasons as follows. If it has, S_2 must have the form:

$$S_2 = (a_1b_1c_1)(d_1e_1f_1)\cdots(a_2)(b_2)\cdots(f_2),$$

and $S_2^2S_1S_2$ unites cycles of S_1, has at most five letters new to S_1, and consequently a single cycle new to S_1. Then

$$S_2^2S_1S_2 = S' = (b_1a_2\cdot)(c_1b_2\cdot)(a_1c_2\cdot)(e_1d_2\cdot)(f_1e_2\cdot)(d_1f_2\cdot)(\alpha_1\alpha_2\alpha_3).$$

The transform of S_1 by S' has no cycle new to S_1, and hence by hypothesis cannot connect cycles of S_1. Now S' fixes the three g's, for if it did not, $\{S_1, S'\}$ would be of degree at most 29, with at most three transitive constituents. It has been proved † that in such a group G, there is always a substitution of order 3 and degree 21, connecting the transitive constituents of $\{S_1, S'\}$ and bringing in at most one new letter to a cycle. Then if we call such a substitution T_1, $\{S_1, S', T_1\}$ is of degree at most 36, and has at most two transitive constituents. Choosing a substitution T_2 connecting the constituents of $\{S_1, S', T_1\}$ and bringing in at most seven new letters, we have a group $\{S_1, S', T_1, T_2\}$ which is transitive and of degree at most 43. Hence,

* Cf. Manning, *Transactions of the American Mathematical Society*, Vol. 20 (1919), No. 1, pp. 73-75.

† Manning, *Transactions of the American Mathematical Society*, Vol. 12 (1911), No. 4, pp. 375-380.

$$S' = (b_1a_2c_3)(c_1b_2a_3)(a_1c_2b_3)(e_1d_2f_3)(f_1e_2d_3)(d_1f_2e_3)(\alpha_1\alpha_2\alpha_3),$$

the $a_3, b_3, c_3, \cdots, f_3$ occurring as they do, so that S' and S_1 are commutative, or otherwise the class of $\{S_1, S'\}$ would be less than 21.

Taking S' as it is given above, we have a group $\{S_1, S'\}$ of degree 24 with four transitive constituents. Picking from G a substitution T_1 connecting transitive constituents of $\{S_1, S'\}$ and bringing in at most one new letter to a cycle, we have the group $\{S_1, S', T_1\}$ of degree at most 31 with at most three transitive constituents. Repeating this twice again, we have the transitive group $\{S_1, S', T_1, T_2, T_3\}$ of degree at most 45. Therefore, $S_1^2S_2S_1$ has not two cycles new to S_2.

Now assume that $S_1^2S_2S_1$ has exactly one cycle new to S_2; it is assumed as before that $S_1^2S_2S_1$ connects cycles of S_2. Therefore:

$$S_2 = (a_1b_1c_1)(d_1\beta_1\cdot)(d_2\beta_2\cdot)(d_3\beta_3\cdot)(y_1y_2y_3)(\cdots)(\cdots)(a_2)(b_2)(c_2),$$

the β's following the d's in order that $S_2S_1S_2^2$ may have a cycle entirely new to S_1.

$$S' = S_2^2S_1S_2 = (b_1a_2\cdot)(c_1b_2\cdot)(a_1c_2\cdot)(\beta_1\beta_2\beta_3)(\cdots)(\cdots)(\cdots),$$

with at most five letters new to S_1. Also, S' must have a cycle new to S_2, but including three letters of S_1. They cannot be letters of two or more cycles of S_1, for $S_2S'S_2^2$ (which is S_1) has this cycle unchanged. It is also a cycle of S_1. Hence it is $(e_1e_2e_3)$, say. Also, a_3, b_3, c_3 cannot follow β_1, β_2, or β_3 in S_2, for

$$S_1^2S'S_1 = \cdots(e_1e_2e_3)(\beta_1\beta_2\beta_3)\cdots,$$

and

$$S_1^2S'S_1S'^2 = (e_1)(e_2)(e_3)(\beta_1)(\beta_2)(\beta_3)\cdots,$$

which is of degree less than 21. Since the class of the group is 21, $S_1^2S'S_1S'^2$ is the identity. We conclude then that S_1 and S' are commutative. Hence, a_3, b_3, and c_3 occur in a single cycle of S_2, so that now

$$S_2 = (a_1b_1c_1)(d_1\beta_1\cdot)(d_2\beta_2\cdot)(d_3\beta_3\cdot)(a_3c_3b_3)(y_1y_2y_3)(\cdots)(a_2)(b_2)(c_2),$$

and

$$S' = (b_1a_2c_3)(c_1b_2a_3)(a_1c_2b_3)(e_1e_2e_3)(\beta_1\beta_2\beta_3)(\cdots)(\cdots).$$

The two empty cycles must be commutative with S_1. If they are filled by powers of any two cycles of [illegible] then a power of S_1, multiplied by S' would be of degree at most 18. T[illegible] cannot be filled by letters new to S_1, because the number of new letters is at most 5. Hence, $S_1^2S_2S_1$ does not connect cycles of S_2.

Now consider

$$S_2 = (a_1b_1\cdot)(y_1y_2y_3)(\cdot\cdot\cdot)\cdot\cdot\cdot,$$

and

$$S_1^2S_2S_1 = (a_2b_2\cdot)(y_1y_2y_3)(\cdot\cdot\cdot)\cdot\cdot\cdot,$$

uniting no two cycles of S_2. Suppose S_2 replaces an a by an a. Then $S_2^2S_1S_2$ fixes y_1, y_2, y_3, connects cycles of S_1, and therefore has a cycle new to S_1, say $(\beta_1\beta_2\beta_3)$. Hence,

$$S_2 = (a_1b_1\cdot)(y_1y_2y_3)(d_1\beta_1\cdot)(d_2\beta_2\cdot)(d_3\beta_3\cdot)(\cdot\cdot\cdot)(\cdot\cdot\cdot).$$

But $S_1^2S_2S_1$ connects the third and fourth cycles of S_2 contrary to hypothesis. Therefore,

$$S_2 = (a_1b_1\cdot)(a_2\cdot\cdot)(a_3\cdot\cdot)(y_1y_2y_3)\cdot\cdot\cdot.$$

Similarly S_2 must displace b_1, b_2, b_3 in three different cycles. Now

$$S_1^2S_2S_1 = (a_2b_2\cdot)(a_3\cdot\cdot)(a_1\cdot\cdot)(y_1y_2y_3)\cdot\cdot\cdot.$$

As b_2 is in a cycle with a_2, then b_3 must be with a_3, and there is no letter new to S_1 in the first three cycles of S_2. For if there were, $S_1^2S_2S_1$ would connect cycles of S_2. Therefore,

$$S_2 = (a_1b_1c_1)(a_2\cdot\cdot)(a_3\cdot\cdot)(y_1y_2y_3)\cdot\cdot\cdot,$$

and the second cycle of S_2 must contain both b_2 and c_2, as

$$S_1^2S_2S_1 = (a_2b_2c_2)(a_3\cdot\cdot)(a_1\cdot\cdot)(\cdot\cdot\cdot)\cdot\cdot\cdot,$$

does not unite cycles of S_2. If $S_2 = (a_1b_1c_1)(a_2c_2b_2)\cdot\cdot\cdot$, then

$$S' = (b_1c_2\cdot)(c_1a_2\cdot)\cdot\cdot\cdot(\gamma_1\gamma_2\gamma_3),$$

making S_2 contain cycles $(d_1\gamma_1\cdot)(d_2\gamma_2\cdot)(d_3\gamma_3\cdot)$ already shown to be impossible. Also $S_1^2S_2S_1 = (a_2b_2c_2)(a_3bc)(a_1bc)\cdot\cdot\cdot$, and as it must not unite cycles of S_2,

$$S_1^2S_2S_1 = (a_1b_1c_1)(a_2b_2c_2)(a_3b_3c_3)\cdot\cdot\cdot,$$

and hence,

$$S_2 = (a_1b_1c_1)(a_2b_2c_2)(a_3b_3c_3)(y_1y_2y_3)\cdot\cdot\cdot.$$

Therefore $S_1^2S_2S_1S_2^2$ is the identity, or is of degree less than 21. But the class of the group is 21, and hence $S_1^2S_2S_1S_2^2$ is the identity; hence $S_1S_2 = S_2S_1$.

We therefore conclude that if two substitutions S_i and S_j of order 3 and degree 21 are of such a form that the first unites cycles of the second and has one or more cycles new to the second, then S_i and S_j are commutative.

We assume now that every substitution of order 3 and degree 21, that

connects cycles of S_1, has four cycles new to S_1. Call this substitution S_2. Since S_2 must be commutative with S_1, it must have the form:

$$S_2 = (a_1b_1c_1)(a_2b_2c_2)(a_3b_3c_3)(\alpha_1\alpha_2\alpha_3)(\beta_1\beta_2\beta_3)(\gamma_1\gamma_2\gamma_3)(\delta_1\delta_2\delta_3).$$

There is an $S_3 = (a_1d_1\cdot)\cdots$, that connects cycles of S_1 and therefore must be commutative with S_1. At first glance, the blank in the first cycle of S_3 may be filled by a b_1, a b_2, a b_3, or an e_1. If it is filled by a b_2 or a b_3, it connects cycles of S_2 and hence must be commutative with S_2. But this is impossible because S_3 has the cycle $(a_1d_1\cdot)$, d_1 a letter new to S_2. If $S_3 = (a_1d_1e_1)\cdots$, being commutative with S_1, it must be of the form:

$$(a_1d_1e_1)(a_2d_2e_2)(a_3d_3e_3)\cdots,$$

with four cycles new to S_1. Then

$$S_3^2S_2S_3 = (d_1b_1c_1)(d_2b_2c_2)(d_3b_3c_3)\cdots,$$

connecting cycles of S_1 and with them four cycles new to S_1. Transforming this by $(a_1c_1)(a_2c_2)(a_3c_3)(\alpha_2\alpha_3)(\beta_2\beta_3)(\gamma_2\gamma_3)(\delta_2\delta_3)$, under which $\{S_1, S_2\}$ is invariant, we have

$$S_3^2S_2S_3 = (a_1d_1b_1)(a_2d_2b_2)(a_3d_3b_3)\cdots,$$

and taking the square of this and calling it S_3, we have:

$$S_3 = (a_1b_1d_1)(a_2b_2d_2)(a_3b_3d_3)\cdots.$$

Hence we need only consider this substitution S_3 commutative with S_1 and connecting cycles of S_1 as

$$S_3 = (a_1b_1d_1)(a_2b_2d_2)(a_3b_3d_3)\cdots,$$

with four cycles new to S_1. These cycles cannot all be new to S_2, for the class of the group is 21. Hence one cycle, say the fourth, has the form $(\alpha_1\cdot\cdot)$. Since the group is doubly transitive, there is a substitution of order 3 and degree 21 of the form: $S_4 = (a_1\alpha_1\cdot)\cdots$, connecting cycles of S_2 and hence commutative with S_2, with four cycles new to S_2. Hence

$$S_4 = (a_1\alpha_1\cdot)(b_1\alpha_2\cdot)(c_1\alpha_3\cdot)(\cdot\cdot\cdot)\cdots,$$

but it is seen that S_4 connects cycles of S_3 and therefore must be commutative with it; hence,

$$S_4 = (a_1\alpha_1\cdot)(b_1\alpha_2\cdot)(c_1\alpha_3\cdot)(d_1\alpha_4\cdot)\cdots,$$

and therefore

$$S_3 = (a_1b_1d_1)(a_2b_2d_2)(a_3b_3d_3)(\alpha_1\alpha_2\alpha_4)\cdots.$$

Similarly, due to the fact that the group is doubly transitive, there is another $S'_4 = (a_1\alpha_2 \cdot) \cdot \cdot \cdot$; but this connects cycles of S_3 and therefore

$$S'_4 = (a_1\alpha_2 \cdot)(b_1\alpha_4 \cdot) \cdot \cdot \cdot,$$

but this is evidently impossible with S_2.

We next assume that every substitution of order 3 and degree 21 that connects cycles of S_1, has at least three cycles new to S_1, and one such has exactly three cycles. Since S_2 and S_1 must be commutative,

$$S_2 = (a_1b_1c_1)(a_2b_2c_2)(a_3b_3c_3)(d_1d_2d_3)(\alpha_1\alpha_2\alpha_3)(\beta_1\beta_2\beta_3)(\gamma_1\gamma_2\gamma_3)$$

or $$S_2 = (a_1b_1c_1)(a_2b_2c_2)(a_3b_3c_3)(d_1d_3d_2)(\alpha_1\alpha_2\alpha_3)(\beta_1\beta_2\beta_3)(\gamma_1\gamma_2\gamma_3).$$

Likewise there is an S_3 commutative with S_1 of the form:

$$S_3 = (a_1d_1 \cdot)(a_2d_2 \cdot)(a_3d_3 \cdot) \cdot \cdot \cdot;$$

but S_3 connects cycles of S_2 and therefore should be commutative with it.

In a similar way we dispose of the assumption that every substitution of degree 21 and order 3, that connects cycles of S_1, has at least two cycles of letters entirely new to S_1, and one such has exactly two cycles.

We now assume that S_2 has only one cycle new to S_1. Since S_1 and S_2 must be commutative, S_2 has the form:

$$S_2 = (a_1b_1c_1)(a_2b_2c_2)(a_3b_3c_3)(d_1e_1f_1)(d_2e_2f_2)(d_3e_3f_3)(\alpha_1\alpha_2\alpha_3).$$

There must also be another substitution S_3 commutative with S_1, $S_3 = (a_1d_1 \cdot) \cdot \cdot \cdot$; but this connects cycles of S_2 and therefore must be commutative with S_2. Hence,

$$S_3 = (a_1d_1 \cdot)(a_2d_2 \cdot)(a_3d_3 \cdot)(b_1e_1 \cdot)(b_2e_2 \cdot)(b_3e_3 \cdot)(c_1f_1 \cdot) \cdot \cdot \cdot,$$

but this is obviously impossible, since S_3 must be of degree 21 and one cycle must be of letters new to both S_1 and S_2.

We are therefore led to the conclusion that for some S_1, there is a substitution S_2 of order 3 and degree 21 that connects cycles of S_1 and has no cycle of letters entirely new to S_1. From those substitutions of degree 21 and order 3 that connect cycles of S_1 and have no cycle of letters new to S_1, we choose one that has the minimum number of letters new to S_1. Call this substitution S_2. Considering the case where S_2 unites only two cycles of S_1, we note that if S_2 has only one cycle in the a's and b's, we can use instead, $S_2^2S_1S_2$, which has two cycles in the a's and b's, and no new letter in these two cycles; for a transitive constituent of degree 7 would lower the class and hence cannot occur.

If S_2 displaces the six a's and b's in just two cycles we can by successive transformations of S_1, use for S_2 a substitution with the six a's and b's in just two cycles, and with at most one new letter in any of the remaining cycles. It remains possible that S_2 has its a's and b's in two cycles with two new letters as, $S_2 = (b_1a_1b)(a_2 \cdot \cdot) \cdot \cdot \cdot$, but by successive transformations of S_1, we can say that S_2 has no more than one new letter in any cycle. Thus we may have a transitive constituent of degree 8, generated by two substitutions of order 3 and degree 6.

If $S_2 = (a_1b_1 \cdot) \cdot \cdot \cdot$, replacing an a by a b or a b by an a, then either $S_2^2S_1S_2$ or $S_2S_1S_2^2$ will have a's and b's in only two cycles and at most only one new letter in any one cycle. We can then call this S_2.

Now it is known that there are no simply transitive primitive groups of degree 8. We now consider an imprimitive constituent of degree 8, generated by two substitutions of order 3 and degree 6, with four systems of imprimitivity of two letters each. S_2 fixes one of the systems of S_1, say α, β and connects cycles of S_1. The other three systems are then (a_1, b_1), (a_2, b_2), and (a_3, b_3). Now $S_1 = (a_1a_2a_3)(b_1b_2b_3) \cdot \cdot \cdot$, and S_2 may take the form, to satisfy the above, of $S_2 = (a_1b_2\alpha)(b_1a_2\beta) \cdot \cdot \cdot$.

Now
$$S_1S_2 = (a_1\beta b_1\alpha)(a_2a_3b_2b_3) \cdot \cdot \cdot,$$
$$S_1S_2 \cdot S_1S_2 = (a_1b_1)(\alpha\beta)(a_2b_2)(a_3b_3) \cdot \cdot \cdot,$$
$$S_1^2S_2 = (a_1a_3\beta b_1b_3\alpha)(a_2b_2) \cdot \cdot \cdot,$$

with the remaining cycles of order 2 or 3. We have then an imprimitive group on the eight letters $a_1, \cdot \cdot \cdot, b_3, \alpha, \beta$ with four systems of two letters each, with S_2 and $S_1^2S_2$ as given and the remaining alternating constituents on three or four letters. If one comes out of order 2, squaring gives six cycles of three letters each, or degree 18. Hence they are all of order 3; but cubing gives a substitution of order 2 and degree 8. Hence such a situation is impossible, since the class is 21.

The other possibility is that we have a transitive constituent of degree 6, in which case
$$S_1 = (a_1a_2a_3)(b_1b_2b_3) \cdot \cdot \cdot,$$
$$S_2 = (a_1b_2a_3)(b_1a_2b_3) \cdot \cdot \cdot.$$

Now there are no simply transitive primitive groups of degree 6. The only possible systems of imprimitivity are three systems of two letters each; say (a_1, b_1), (a_2, b_2), and (a_3, b_3), and the group in the systems is of order 3. Now
$$S_1S_2 = (a_1b_3a_2)(a_3b_2b_1) \cdot \cdot \cdot,$$
$$S_1^2S_2 = (a_1)(b_1)(a_2b_2)(a_3b_3) \cdot \cdot \cdot.$$
Also,
$$S_1S_2S_1 = (a_3b_3)(a_1b_1) \cdot \cdot \cdot,$$
and
$$S_2S_1^2 = (a_1b_1)(a_2b_2) \cdot \cdot \cdot.$$

Thus S_2 with S_1 generates a group of order 12 and degree 6 in $a_1, \cdots, b_3$.

Suppose S_2 has $a_1, a_2, a_3, b_1, b_2, b_3$ in exactly three cycles. The group on $a_1, \cdots, b_3$ is a primitive group of class 6, or an alternating group, but the latter is impossible, due to the presence of substitutions of order 7. If it is of class 6, it is 2-ply transitive so that $\{S_1, S_2\}$ contains a transform of S_1, equal to $(a_1 b_1 \cdot) \cdots$ of degree 6, but this has already been considered.

Suppose S_2 has $a_1, \cdots, b_3$ in exactly four cycles. If

$$S_2 = (a_1 b_1 \cdot)(\cdots)(\cdots)(\cdots)(e_1 f_1 g_1) \cdots (e_2)(f_2)(g_2),$$

$\{S_1, S_2\}$ has at most four transitive constituents and is of degree at most 31, leading to a transitive group of degree at most 52. Hence

$$S_1^2 S_2 S_1 = (a_2 b_2 \alpha_1)(\alpha_2 \cdot \cdot)(\alpha_3 \alpha_5 \cdot)(\alpha_4 \alpha_6 \cdot) \cdots$$

has at most six letters new to S_2; therefore $\{S_2, S_1^2 S_2 S_1, S_1 S_2 S_1^2\}$ has one transitive constituent $a_1, \cdots, \alpha_6$ of degree 12, and has at most four transitive constituents and a degree of at most 27, leading to a transitive group G of degree at most 48.

If $S_2 = (a_1 b_1 \alpha) \cdots$ and has the a's and b's in five cycles, a repetition of the above shows that the degree of the possible group is well within our limit. Hence, $H_2 = \{S_1, S_2\}$ is a group of degree at most 26, containing at most six transitive constituents.

We also show that H_2 has at most one transitive constituent of degree 3. We have:

$$S_1 = (a_1 a_2 a_3)(b_1 b_2 b_3)(c_1 c_2 c_3) \cdots,$$

and

$$S_2 = (a_1 b_2 a_3)(b_1 a_2 b_3) \cdots;$$

then

$$S_1 S_2 = (a_1 b_3 a_2)(a_3 b_2 b_1) \cdots.$$

If H_2 has two transitive constituents of degree 3,

$$S_1 S_2 = \cdots (c_1 c_3 c_2)(d_1 d_3 d_2) \cdots,$$

and then

$$S_2 = \cdots (c_1 c_2 c_3)(d_1 d_2 d_3) \cdots.$$

So then $S_1^2 S_2 = (a_2 b_2)(a_3 b_3)(a_1)(b_1) \cdots$ fixes the six letters $c_1, \cdots, d_3$ and hence is of degree at most 20.

We can also show that from $H_2 = \{S_1, S_2\}$, we can build up a transitive group of degree at most 53. We take for S_3, $S_3 = (a_1 d_1 \cdot) \cdots$. Now $S_3^2 S_1 S_3$ replaces an a (or a b) by a d, unless S_3 has three a's (or three b's) in three different cycles; $S_3^2 S_2 S_3$ does the same unless b_2 is in a fourth cycle of S_3; $S_3^2 S_1 S_2 S_1 S_3$ does also, unless b_3 is in a fifth cycle; and similarly for b_1 in a sixth cycle. But S_3 cannot have $a_1, \cdots, b_3$ in more than five cycles. This all means then that for S_3 a transform of a substitution of

order 3 of H_2 may be used, and it has the letters $a_1, \cdots, b_3, d_1, d_2, d_3$ in exactly three cycles. Then the transitive constituent $a_1 \cdots$ of $\{H_2, S_3\}$ is of degree 11, 12, or 13, if S_3 brings in more than five new letters. Constituents of degree 11 or 13 are impossible. Since all primitive groups of degree 12 are 2-ply transitive and therefore contain a subgroup of degree 11, which would lower the class, we conclude that this constituent is imprimitive. Since it is generated by substitutions of order 3 and degree 9, it can have only systems of three letters each, and therefore four in number. Suppose S_2 has a letter d_4; S_1 fixes the system to which d_4 belongs, and therefore S_2 permutes it, which is impossible.

Since we have shown that H_2 has at most one transitive constituent of degree 3, or in other words that all but one of the transitive constituents of $H_3 = \{H_2, S_3\}$, other than $a_1, \cdots, b_3$, are of degree 4, we conclude that of the substitutions

$$\begin{gathered}(a_1c_1 \cdot) \cdots \\ (a_1d_1 \cdot) \cdots \\ (a_1e_1 \cdot) \cdots \\ (a_1f_1 \cdot) \cdots \\ (a_1g_1 \cdot) \cdots,\end{gathered}$$

only one can displace more than five letters new to H_2. Hence H_2 leads to a transitive group of degree at most 53.

There exists a substitution $S_3 = (a_1g_1 \cdot) \cdots$, and we say that among all the substitutions of order 3 and degree 21, which replace a letter of the transitive constituent $a_1, \cdots, b_3$ by a letter of another constituent of H_2, S_3 displaces a minimum number of letters new to H_2. Since the constituent $a_1, \cdots$ of H_2 is of degree less than 8, we cannot be sure that S_3 has at most one letter new to H_2 in a cycle. This will be true if S_3 replaces one of the six letters $a_1, \cdots, b_3$ by one of them, or if it replaces one of the letters $g_1 \cdots$ (there are three or four of them) by a g; and what is true of $g_1, \cdots$ is true of any other transitive constituent $f_1, \cdots$, with letters in the cycles of S_3 with $a_1, \cdots, b_3$. Suppose that S_3 has two new letters in some one cycle. Then

$$S_3 = (a_1g_1 \cdot)(a_2 \cdot \cdot)(a_3 \cdot \cdot)(b_1 \cdot \cdot)(b_2 \cdot \cdot)(b_3 \cdot \cdot) \cdots.$$

Now S_3 has two new letters adjacent, so that $S_3H_2S_3^{\pm 1}$ fixes a new letter. Hence $S_1^2S_3S_1$ displaces the same eighteen letters as S_3 in its first six cycles.

We next show that $S_1^2S_3S_1$ has no cycle new to S_3. If so,

$$S_3 = (d_1e_1f_1) \cdots (d_2)(e_2)(f_2),$$

so that

$$S_1^2S_3S_1 = (d_2e_2f_2) \cdots.$$

But now $\{S_1, S_3\}$ has one or two transitive constituents in $a_1, \cdots, b_3, g_1, \cdots,$ one in $d_1, \cdots, e_1, \cdots, f_1, \cdots,$ and one in the c's; that is, at most four transitive constituents and is of degree $\leqq 30$, leading to a transitive group of degree at most 51. Hence neither $S_1{}^2S_3S_1$ nor $S_1S_3S_1{}^2$ has a cycle new to S_3. Then $\{S_3, S_1{}^2S_3S_1\}$ is of degree at most 23, and $\{S_3, S_1{}^2S_3S_1, S_1S_3S_1{}^2\}$ is of degree at most 25.

If S_3 has a letter new to S_1 with all the a's, and also a letter new to S_1 with the b's, then $\{S_3, S_1{}^2S_3S_1, S_1S_3S_1{}^2\}$ has at most three transitive constituents, leading to a transitive group of degree at most 39. Therefore S_3 unites at least two sets of H_2 to $a_1, \cdots, b_3$.

Suppose the last cycle of S_3 is made up of letters new to S_1, say $S_3 = \cdots (\alpha_1\alpha_2\alpha_3)$. If these are the only new letters in S_3, $\{S_1, S_3\}$ is of degree at most 24 and has fewer than four transitive constituents, thus leading to a transitive group of degree at most 38. But if there are other new letters in S_3, $\{S_3, S_1{}^2S_3S_1, S_1S_3S_1{}^2\}$ is of degree 21, has at most five transitive constituents, one of which is of degree 9, leading to a transitive group of degree at most 49.

Now S_3 has no cycle of new letters and does not have new letters with both the a's and the b's. Therefore $\{H_2, S_3\}$ has at most four transitive constituents and is of degree 29 at most, leading to a transitive group of degree at most 50.

Then S_3 has only one new letter in any cycle, and does not unite more than two transitive constituents of H_2. If it united more than two, it would lead to a transitive group of degree at most 54.

Now suppose S_2 connects three cycles of S_1, say the a's, the b's, and the c's. If there is only one new letter in a cycle, then $\{S_1, S_2\}$ is of degree at most 28 with at most five sets of transitivity, leading to a transitive group of degree 56. Hence we are interested only in the case that S_2 has two letters in a cycle new to S_1. We say that S_2 has a minimum number of letters new to S_1.

Suppose $S_2 = (a_1b_1c_1) \cdots$; that is, has only one cycle containing any of the letters $a_1, \cdots, c_3$. Then

$$S'_2 = S_2{}^2S_1S_2 = (b_1a_2a_3)(c_1b_2b_3)(a_1c_2c_3) \cdots$$

connects the three cycles of S_1, and has fewer new letters. But S'_2 might be of the form:

$$S'_2 = (b_1a_2a_3)(c_1b_2b_3)(a_1c_2c_3)(\alpha_1\alpha_2\alpha_3) \cdots,$$

and if so then

$$S'_2{}^2S_1S'_2 = (c_2a_3b_1)(a_2b_3c_1)(b_2c_3a_1) \cdots (\alpha_1)(\alpha_2)(\alpha_3) \cdots.$$

If $S'^2_2S_1S'_2$ has a cycle $(\beta_1\beta_2\beta_3)$, then $\{S'_2, S_1^2S'_2S_1\}$ is of degree at most 24 and has at most three transitive constituents, leading to a transitive group of degree $\leqq 38$. Hence we need not consider the case $S'^2_2S_1S'_2 = \cdots (\beta_1\beta_2\beta_3) \cdots$, or what is the same thing, $S'_2 \neq \cdots (\alpha_1\alpha_2\alpha_3) \cdots$. Hence if S_2 has an a and a b and a c in just one cycle, we have a group of degree well within our limit.

Suppose $S_2 = (a_1b_1 \cdot) \cdots$, with a's, b's, and c's in just two cycles. The question is: Can

$$S_2 = (a_1b_1 \cdot)(\cdots)(d_1\alpha_1 \cdot)(d_2\alpha_2 \cdot)(d_3\alpha_3 \cdot) \cdots ?$$

In $\{S_1, S_2\}$ the transitive constituent in $a_1, \cdots, c_3$ is primitive of class 3 or 6, and therefore 2-ply transitive. Therefore, the transform of S_2 by S_1 unites the first two cycles of S_2 and is of degree 6 in $a_1, \cdots$. But this case has already been considered.

Suppose S_2 has a's, b's, and c's in exactly three cycles. The transitive constituent $a_1, \cdots$ of H_2 is not of degree 11 or 13. The possible degrees are 9, 10, or 12. Suppose first that the transitive constituent $a_1, \cdots$ is imprimitive. There are three letters in each system, and hence three or four systems. Consider the possibility of degree 9. S_2 does not connect any two of the last four cycles of S_1, for if it did, we would have a transitive group of degree at most 53. Hence at least one transitive constituent of H_2 is the alternating group of order 60 and degree 5. Then S_1S_2 has a cycle of order 5. An imprimitive group with less than five systems of three letters each can contain no substitutions of order 5. Therefore degree 9 is out. If this primitive constituent is of degree 12, one system is composed of three new letters in the first three cycles of S_2, so that again S_1S_2 has a cycle of five letters. Therefore, the constituent $a_1, \cdots$ must be primitive, and of degree 9 or 10. (The primitive groups of degree 12 are 2-ply transitive.) As we saw above, it must contain a substitution of order 5. A positive primitive group of degree 9, in which there is a substitution of degree and order 5, is the alternating group and hence it brings in substitutions of order 7 which are impossible. Therefore, $a_1, \cdots$ is a primitive group of degree 10, if any. Hence $S_2^2S_1S_2$ certainly fixes a new letter, has no cycle new to S_1, and connects cycles of S_1.

Suppose S_2 has the a's, b's, and c's in exactly four cycles. Now S_2 has at most twelve new letters, and hence S_2 connects no other cycles of S_1. If

$$S_2 = (a_1b_1 \cdot)(\cdots)(\cdots)(\cdots)(d_1\alpha_1 \cdot)(d_2\alpha_2 \cdot)(d_3\alpha_3 \cdot).$$

it fixes the nine letters $e_1, \cdots, f_1, \cdots, g_3$ so that

$$S_2^2S_1S_2 = (\cdots)(\cdots)(\cdots)(\alpha_1\alpha_2\alpha_3)(e_1e_2e_3)(f_1f_2f_3)(g_1g_2g_3)$$

and $$S_1^2S_2^2S_1S_2=(d_1d_3d_2)(\alpha_1\alpha_2\alpha_3)\cdots,$$
with five more such cycles. Now S_2 has at most six new letters in the first four cycles; Hence $S_2^2S_1S_2$ has at most four new letters in the first three cycles. Therefore $S_1^2S_2^2S_1S_2$ is of degree at most 19. But the class is 21, so this is impossible.

But now perhaps $S_2=(a_1b_1\alpha_1)(b_2\alpha_2\cdot)(b_3\alpha_3\cdot)\cdots$. Then
$$S_2^2S_1S_2=(b_1\cdot\cdot)(\alpha_1\alpha_2\alpha_3)\cdots.$$
If $S_2^2S_1S_2$ connects all three cycles $a\cdots, b\cdots, c\cdots$, it fixes some letters $a_1,\cdots, b_1,\cdots, c_3$, and $(S_2^2S_1S_2)^2S_1(S_2^2S_1S_2)$ connects two cycles, fixes α_1, and has no new cycle. Hence $S_2^2S_1S_2$ connects just two cycles of S_1, and this case has already been considered.

Hence S_2 has the three a's in three cycles, the three b's in three cycles, and the three c's in three cycles. So say that
$$S_2=(a_1b_1\cdot)(a_2\cdot\cdot)(a_3\cdot\cdot)(\cdots)\cdots.$$
The constituent $a_1,\cdots$ is of degree 12, being generated by two substitutions of order 3, one of three cycles and one of four cycles, and cannot be primitive or a substitution of order 11 comes in. It cannot have systems of imprimitivity of two letters each, because of its two generators; nor can it have systems of six letters each for the same reason. Hence the only possibilities are that there are three systems of four letters each, or four systems of three letters each. Suppose there are three new letters a_4, b_4, c_4 in the $a_1,\cdots$ set. S_1 fixes systems to which a_4 belongs, so either (a_4, b_4, c_4) is a system or (a_1, a_2, a_3, a_4) is a system (or both). If the former is a system, (a_4, b_4, c_4) are in three cycles of S_2. If the systems are (a_1, a_2, a_3, a_4), (b_1, b_2, b_3, b_4), and (c_1, c_2, c_3, c_4),
$$S_2=(a_1b_1c)(a_2bc)(a_3bc)(a_4bc)\cdots,$$
and there are at least five new letters in the last three cycles, and S_2 can be written as
$$S_2=\cdots(d_1\alpha_1\alpha_2)(\cdot\alpha_3\alpha_4)(\cdot\cdot\alpha_5)(\cdots).$$
Suppose S_2 fixes d_2 and d_3; then $S_1S_2=(d_1d_2d_3\alpha_1\alpha_2)\cdots$, or a cycle of order 5, which is impossible. If S_2 fixes d_3 only, then
$$S_2=\cdots(d_1\alpha_1\alpha_2)(d_2\alpha_3\alpha_4)\cdots,$$
and S_1S_2 has a cycle of order 7, also impossible. Suppose
$$S_2\neq\cdots(d_1\alpha_1\alpha_2)(d_2\alpha_3\alpha_4)(d_3\alpha_5\alpha_6)\cdots,$$
as we have already seen. Then
$$S_1^2S_2^2S_1S_2=(d_1d_3d_2)(\alpha_1\alpha_2\alpha_3)\cdots,$$

and is of degree $\leqq 18$.

If $$S_2 = \cdots (d_1\alpha_1\alpha_2)(d_2d_3\alpha_3)\cdots,$$
or if $$S_2 = \cdots (d_1\alpha_1\alpha_2)(d_3d_2\alpha_3)\cdots,$$

both contain a cycle of the form $(e_1\beta_1\beta_2)$, but this is the same as the case where S_2 fixes d_2 and d_3, only now it is in the e's.

Suppose the a's, b's, and c's are strung out in exactly five cycles. We have:

$$S_2 = (a_1b_1\cdot)(\cdots)(\cdots)(\cdots)(\cdots)\cdots,$$

with at least one of the nine letters $a_1, \cdots, c_3$ in each of the first five cycles. Then at least two sets of S_1, say the f's and the g's, do not appear in S_2. There will be a cycle in $S_2^2S_1S_2$ new to S_1, only if

$$S_2 = (a_1b_1\alpha_1)(b_2\alpha_2\cdot)(b_3\alpha_3\cdot)\cdots,$$
but then $$S_2^2S_1S_2 = (b_1\cdot\cdot)(\alpha_1\alpha_2\alpha_3)(\cdots)\cdots,$$

would connect at least three cycles by letters scattered in at most three cycles, a case already disposed of. Hence there are the nine letters $a_1, \cdots, c_3$ in S_2, and then $\{S_2, S_1^2S_2S_1, S_1S_2S_1^2\}$ is of degree at most 25 with at most four transitive constituents, leading to a transitive group of degree at most 46.

Suppose that just six cycles contain at least one of the first nine letters of S_1. Then

$$S_2 = (a_1b_1\cdot)(\cdots)(\cdots)(\cdots)(\cdots)(\cdots)\cdots,$$

with not more than one cycle containing three of the nine letters, and with not more than one of the nine letters fixed. Then the group $\{S_2, S_1^2S_2S_1, S_1S_2S_1^2\}$ is of degree at most 24 with at most three transitive constituents, leading to a transitive group of degree at most 38.

If the nine letters $a_1, \cdots, c_3$ are in seven cycles, all nine occur, and not more than one cycle can displace three of the nine letters. In all cases there is a new letter in at least one cycle containing an a, in at least one containing a b, and in at least one containing a c. Then the group $\{S_2, S_1^2S_2S_1, S_1S_2S_1^2\}$ is of degree at most 21, and is transitive.

We arrive then at the conclusion that if a doubly transitive group G contains a substitution S_1 of order 3 and degree 21, it contains a transitive subgroup H of degree at most 56. We saw that if H is primitive, the degree of G is at most 57; we have yet to show that if H is imprimitive, the degree of G is at most 57.

We take the most unfavorable case, when the degree of H is 56 and H has eight systems of imprimitivity of seven letters each. If we can prove that there is only one possible system of imprimitivity to which a given letter α_1

can belong,* then we know that the degree of G is at most 57. H is generated by six substitutions $S_1, S_2, \cdots, S_6$, for S_2 connects three systems of transitivity of S_1. Now S_6 connects the set $a_1, \cdots$ to the only remaining set of transitivity of H_5, say the $g_1, \cdots$ set, and also brings in seven new letters, by hypothesis. Therefore,

$$S_6 = (a_1 g_1 \alpha_1)(\cdot\cdot\alpha_2)(\cdot\cdot\alpha_3)\cdots(\cdot\cdot\alpha_7).$$

It is apparent that α_1 and a_1 never belong to the same system. Moreover, no letter $a_1, \cdots$ of H_5 is in a system with α_1, because the transforms of S_6 by the substitutions of H_5 all have a first cycle of the form $(a_i g_i \alpha_1)$, and what is true for α_1 and a_1 is true for α_1 and a_i, where a_i represents any one of the set $a_1, \cdots$. The same is true for α_1 in respect to the members of the set $g_1, \cdots$. Therefore, no letter of $a_1, \cdots, g_1, \cdots$ is in a system with α_1. Hence, α_1 can be in only one system, namely $\alpha_1, \cdots, \alpha_7$.

If H is of degree 54, there may be nine systems of imprimitivity of six letters each, but in exactly the same way it can be proved that a given letter α_1 belongs to only one possible system, namely $\alpha_1, \alpha_2, \cdots, \alpha_6$. Hence G is of degree at most 55.

Less difficult means dispose of the restrictions upon n for $\epsilon = 2, 3$.

9. Theorem II. *The class u (> 3) of a triply transitive group of degree n, that contains a substitution of degree $u + \epsilon$, ϵ a positive integer, and of order p^c (p an odd prime), is greater than*

$$(n/2)(1 - 1/p^c) - 4\epsilon, \quad \textit{provided} \quad \epsilon < n/30.$$

10. In order to deal successfully with this problem, we must establish a lemma, in the proof of which we shall need Bochert's Lemma: †

If the substitutions S and T have exactly m letters in common and if S replaces q, and T r, common letters by common letters, the degree of $S^{-1}T^{-1}ST$ is not greater than $3m - q - r$.

Lemma: *If S and T have two regular substitutions of degree $u + \epsilon$ and odd order d, which generate a group of class u, and if no power of T is commutative with a power of S, S and T have at least*

$$u/2 - u/2d - [(d-1)(d-3)/2d(d+3)]\epsilon$$

letters in common.

If S and T have m common letters, Bochert's Lemma tells us that m is

* Manning, *Primitive Groups*, 1921, p. 93.

† Bochert, *Mathematische Annalen*, Vol. 40 (1892), p. 176.

at least the integral part of $u/3$. We assume $d \geqq 5$. Say * S has s_i cycles, each of which contains i letters in common with T, and T has t_j cycles, each of which contains j letters in common with S; then we have

(A) $$s_1 + s_2 + \cdots + s_d = t_1 + t_2 + \cdots + t_d = (u + \epsilon)/d$$

also

(B) $$s_1 + 2s_2 + \cdots + dS_d = t_1 + 2t_2 + \cdots + dt_d = u/3 - \xi + k_1$$
$$= u/3 + k,$$

where $\xi = 0$, $1/3$, or $2/3$ so that $u/3 - \xi$ is an integer, and k_1 is a positive integer or zero.

For the average number of sequences of common letters in the $d-1$ powers of the substitution S we have

$$[2s_2 + 6s_3 + \cdots + i(i-1)s_i + \cdots + d(d-1)s_d]/(d-1).$$

Take S^v and T^w with the only restriction that each must contain at least the average number of sequences of common letters. Bochert's Lemma limits the degree of $S^{-v}T^{-w}S^vT^w$ to $3m - q - r = u + 3k - q - r$, from which we conclude that the total number of sequences of common letters in S^v and T^w cannot exceed $3k$, since S^v and T^w are non-commutative and the degree of their commutator is at least u. It follows then that

$$\sum_0^d [i(i-1)/(d-1)]s_i + \sum_0^d [j(j-1)/(d-1)]t_j \leqq 3k,$$

and one of these summations is at most $3k/2$; hence

(C) $$\sum_0^d [i(i-1)/(d-1)]\, s_i \leqq 3k/2.$$

If from the three equations (A), (B), and (C), s_x and s_y are eliminated $(y > x)$,

$$u \leqq \frac{3d}{2x}\left(\frac{3d-4x-3}{2d-3x-3}\right)k + \frac{3(x+1)}{2d-3x-3}\,\epsilon.$$

We determine a value of x (a positive integer less than d) that will make the coefficient of k a minimum. The minimum is for

$$x = \frac{3(d-1) - [(d+2)(d-1)]^{1/2}}{4}\,.$$

For $x = (d-3)/2$ and $x = (d-1)/2$ the coefficient of k has the same value, but for the former the coefficient of ϵ is the smaller. Hence

$$u \leqq [6d/(d-3)]k + [3(d-1)/(d+3)]\epsilon.$$

---

* Cf. Manning, *Transactions of the American Mathematical Society*, Vol. 31 (1929), No. 4, pp. 644 ff.

But $m \geqq u/3 + k$; hence

$$m \geqq u/2 - u/2d - [(d-1)(d-3)/2d(d+3)]\epsilon,$$

which proves the lemma.

11. By hypothesis there is in G a substitution S of degree $u+\epsilon$ and order p^c. It has the form: $S = (a \cdots b \cdots) \cdots$. Since G is triply transitive, there exists a substitution S' similar to S that displaces a and fixes b. Transform S' by the w substitutions of a subgroup that fixes a and b. No power of $S^{(i)}$ is commutative with S. The w conjugates displace exactly

$$w + w(u+\epsilon-1)(u+\epsilon-2)/(n-2)$$

letters of S. Our lemma states that when S and $S^{(i)}$ are regular substitutions, the degree of $\{S, S^{(i)}\}$ is limited. Over the complete set of w substitutions

$$\begin{aligned} w + w(u+\epsilon-1)(u+\epsilon-2)/(n-2) \\ \geqq w\{u/2 - u/2p^c - [(p^c-1)(p^c-3)/2p^c(p^c+3)]\epsilon\}, \end{aligned}$$

from which

$$\text{(D)} \quad \frac{n-u}{p^c} - (u-2)\left(\frac{n}{2} - \frac{n}{2p^c} - u - 2\epsilon\right) + \epsilon^2 + \epsilon + (n-2)\frac{(p^c-1)(p^c-3)}{2p^c(p^c+3)}\epsilon \geqq 0.$$

As in the previous case, assume a relation

$$\text{(E)} \qquad u > n/2 - n/2p^c - 2\epsilon - \delta,$$

with δ to be determined so that $u = n/2 - n/2p^c - 2\epsilon - \delta$ fails to satisfy (D). If we put $\delta = 2\epsilon$, (E) is true provided $n > 30\epsilon$, ϵ a positive integer.

12. Theorem III. *If u is the class of a doubly transitive group of degree n, in which there is a substitution of degree $u+\epsilon$ (ϵ a positive integer) and of prime order p,*

$$u > \frac{n}{2}\left(1 - \frac{1}{p}\right) - \frac{n^{1/2}}{2}\left(1 - \frac{1}{p^2}\right)^{1/2} - 4\epsilon, \quad \textit{provided} \quad n \geqq 45\epsilon.$$

Consider a substitution $S = (ab \cdots) \cdots$ of prime order p (>3). Since the group is at least doubly transitive, S is one of a set of w conjugates. Since S is of degree $u+\epsilon$, it contains $u+\epsilon$ sequences of two letters. In the set of w conjugates, any particular sequence in S occurs $w(u+\epsilon)/n(n-1)$ times, and sequences containing one letter of S and a letter new to S occur exactly $2w(u+\epsilon)^2(n-u-\epsilon)/n(n-1)$ times.

The y substitutions $S', S'', \cdots, S^{(y)}$ not commutative with S contain at least one sequence of a letter of S and at letter new to S. If we assume

that when one such sequence occurs in a substitution, $(p-1)(u+\epsilon)/p$ occur in the substitution, then

$$y \geqq 2w(u+\epsilon)(n-u-\epsilon)/n(n-1)(1-1/p).$$

We have again

$$\sum_{w} [x_i-(u+\epsilon)^2/n]^2 = w(u+\epsilon)^2(n-u-\epsilon)^2/n^2(n-1);$$

from our lemma,

$$\sum_{y} x_i \geqq y\{(u/2)(1-1/p)-[(p-1)(p-3)/2p(p+3)]\epsilon\},$$

also

$$\sum_{y} [x_i-(u+\epsilon)^2/n]$$
$$\geqq y\{(u/2)(1-1/p)-[(p-1)(p-3)/2p(p+3)]\epsilon-(u+\epsilon)^2/n\}^2;$$

and since

$$\sum_{w-y} [x_i-(u+\epsilon)^2/n]^2 \geqq [1/(w-y)]\{\sum_{w-y}[x_i-(u+\epsilon)^2/n]\}^2,$$

we have

(A) $(u+\epsilon)(n-u-\epsilon)(1-1/p)/2n-(u+\epsilon)^2(n-u-\epsilon)^2/n^2(n-1)$
$\quad -\{(u/2)(1-1/p)-[(p-1)(p-3)/2p(p+3)]\epsilon-(u+\epsilon)^2/n\}^2 \geqq 0.$

Let

(B) $$u=(n/2)(1-1/p)-(\gamma/2)n^{1/2}-\delta.$$

We choose γ and δ so that (B) fails to satisfy (A). We can then say that

$$u>(n/2)(1-1/p)-(\gamma/2)n^{1/2}-\delta$$

is a true relation. After simplification (A) becomes a polynomial in $n^{1/2}$. We choose $\gamma^2=(1-1/p^2)$ so as to eliminate the highest power of n. δ is so chosen that the next highest power of n is negative. If $\delta=3\epsilon+1$, this is true. Hence for n sufficiently large,

$$u>(n/2)(1-1/p)-(n^{1/2}/2)(1-1/p^2)^{1/2}-3\epsilon-1.$$

We shall, however, for sharper definition of n, take $\delta=4\epsilon$. It is possible to show that

$$u>(n/2)(1-1/p)-(n^{1/2}/2)(1-1/p^2)^{1/2}-4\epsilon, \quad \text{provided} \quad n \geqq 45\epsilon.$$

STANFORD UNIVERSITY,
JULY, 1932.

REPRESENTATION BY EXTENDED POLYGONAL NUMBERS AND BY GENERALIZED POLYGONAL NUMBERS.

By L. W. Griffiths.

1. *Introduction.* Complete results on representation of positive integers by extended polygonal numbers are obtained in this paper; they are similar to my results on representation by polygonal numbers * and to the results on representation by squares proved by Dickson.† This similarity is also true of the results on representation by generalized polygonal numbers, obtained in this paper. These facts are evident from the following definitions and summary of results on representation.

In representation by squares the summands are values of x^2 for $x = 0, 1, 2, \cdots$. In representation by polygonal numbers the summands are polygonal numbers of order $m + 2$, that is, values of

$$(1) \qquad p(x) = x + m(x^2 - 2)/2 \quad \text{for} \quad x = 0, 1, 2, \cdots,$$

with m a fixed positive integer. In representation by extended polygonal numbers the summands are values of

$$(2) \qquad e(x) = -x + m(x^2 + x)/2 \quad \text{for} \quad x = -1, 0, 1, 2, \cdots,$$

with m a fixed positive integer. In representation by generalized polygonal numbers the summands are values of

$$(3) \qquad g(x) = x + m(x^2 - x)/2 \quad \text{for} \quad x = 0, \pm 1, \pm 2, \cdots$$

with m a fixed positive integer. Each of the sets (1), (2), (3) consists of 0, 1, and infinitely many distinct positive integers > 1. If $m = 2$ each is precisely the set of squares. Hence representation by polygonal numbers is a generalization of representation by squares; so also is representation by extended polygonal numbers, and representation by generalized polygonal numbers. If $m = 1$ the sets are identical, being the triangular numbers. If $m \geqq 3$ the sets are distinct, and (3) consists of (1) and (2).

On representation by squares there is the classic theorem that every

* *Annals of Mathematics*, Ser. 2, Vol. 31 (1930), pp. 1-12. This paper will be cited as R. P. N.

† *Bulletin of the American Mathematical Society*, Vol. 33 (1927), pp. 63-70. If the number n of variables is not greater than 3 there are no universal forms, if $n = 5$ there are six, while if $n > 5$ there are no non-trivial ones.

positive integer is a sum of four squares. The statement of this theorem that has been suggestive for generalization is that every positive integer is represented by the form $x^2 + y^2 + z^2 + w^2$, that is, that this form is universal. Again $x^2 + y^2 + z^2 + w^2$ and $x^2 + y^2 + 2z^2 + 2w^2$ are universal forms; more generally, Dickson proved that there are exactly fifty-four universal principal quaternary quadratic forms.

On representation by triangular numbers Liouville * proved that the functions $p_1 + p_2 + p_3$, $p_1 + p_2 + 2p_3$, etc., are universal, that is, that every positive integer is a sum of three triangular numbers, a sum of two triangular numbers and twice a third, etc.

On representation by (1), (2), (3), with $m \geqq 3$, the similarity of results to those on representation by squares is remarkable. Cauchy † was the first to publish a proof of the Fermat theorem that every positive integer A is a sum of $m + 2$ polygonal numbers of order $m + 2$. That is, if m is a fixed but arbitrary integer $\geqq 3$ and A is a positive integer, there are values $p_1, \cdots, p_{m+2}$ of (1) such that $A = p_1 + \cdots + p_{m+2}$; in other words, for (1) the function $p_1 + \cdots + p_{m+2}$ is universal. The similar theorem for extended polygonal numbers of order $m + 2$, namely that every positive integer A is a sum of m values of (2) if $m \geqq 6$, but of $m + 1$ values if $m = 3, 4, 5$, is due to Dickson. ‡ The similar theorem for generalized polygonal numbers, also due to Dickson § states that A is a sum of $m - 2$, $m - 1$, or m values of (3), according as $m \geqq 6$, $m = 5, 4$, or $m = 3$. In these theorems for $m \geqq 3$ the universal functions have all coefficients unity, as in the theorem that $x^2 + y^2 + z^2 + w^2$ is universal.

The determination of all universal functions, not merely those in which every coefficient is unity, is the problem of my papers on representation by (1), (2), (3), with $m \geqq 3$. Its importance is first in being the general case with respect to the coefficients, and second in that the universality of a function having not all coefficients unity implies the universality of other

* *Journal de Mathématiques*, Ser. 2, Vol. 7 (1862), p. 407, and Vol. 8 (1863), p. 73. He found seven universal functions for $n = 3$, none for $n = 1, 2$, but did not consider $n > 3$.

† *Œuvres*, Ser. 2, Vol. 6, pp. 320-353. This includes $m = 1, 2$.

‡ *American Journal of Mathematics*, Vol. 50 (1928), pp. 1-48. This and the Fermat theorem are particular cases in a comprehensive discussion of representation using as summands any quadratic function $q(x)$ which takes integral values $\geqq 0$ for every integer $x \geqq 0$. Note that the definition (2) of the extended polygonal numbers differs from the definition of Dickson in *Bulletin of the American Mathematical Society*, Vol. 34 (1928), p. 205, in including $e(-1) = 1$; this inclusion seemed wise, since no other value of (2) is unity when $m > 2$.

§ *Journal de Mathématiques*, Ser. 9, Vol. 7 (1928), Theorems 11-15.

functions. For example, in the Fermat theorem $m+2$ polygonal numbers are sufficient, that is, in representation by (2) the function having precisely $m+2$ coefficients, each unity, is universal; its universality is implied by that of the function having its first m coefficients unity, the next and last being two. The universality of this latter function is proved in my earlier paper. More generally, in my earlier paper on representation by (1), the problem to find every universal function with the sum of its coefficients $\leqq m+2$ is completely solved. In this present paper on representation by (2) and by (3) a similar condition on the sum of the coefficients is imposed, and for a similar reason. A discussion of representation with no condition on the sum of the coefficients will be given in a later paper.

For representation by (2), all universal functions, with the sum of the coefficients so conditioned, are obtained in this paper. Theorem 1 gives necessary conditions. Theorem 2 proves these conditions sufficient, except perhaps for certain stated integers, relatively small and relatively few. The necessary but extremely arduous direct verification for these integers was not undertaken, since experience indicated that actual verification was practically certain.

For representation by (3), necessary conditions are given by (11). The new lemmas 3 and 6 are vital in the proof that those functions among (11) which are in Theorem 3 are indeed universal, except perhaps again for certain stated integers. Finally, for the functions (11) not included in Theorem 3 there is no conclusion, since it has been impossible to prove for these functions lemmas analogous to lemmas 3 and 6.

2. *Necessary conditions for universality in representation by extended polygonal numbers.* We use (2) and the notations

$$(4) \qquad \begin{aligned} &f = a_1e_1 + \cdots + a_ne_n = (a_1, \cdots, a_n), \quad 1 \leqq a_1 \leqq \cdots \leqq a_n, \\ &w = a_1 + \cdots + a_n, \quad f_k = (a_{k+1}, \cdots, a_n) \quad (1 \leqq k < n). \end{aligned}$$

The proofs of the fundamental lemmas 1, 2 of R. P. N. will hold here if $a_k \leqq m-1$ for every coefficient, and hence if $w \leqq m-1$. It is sufficient to assume, however, merely that $w \leqq m$ for $m \geqq 6$ and $w \leqq m+1$ for $m = 3, 4, 5$. This latter hypothesis is suggested by Dickson's theorem, quoted in § 1, that f is universal if $a_1 = \cdots = a_n = 1$, and $n = m$ for $m \geqq 6$ but $n = m+1$ for $m = 3, 4, 5$. The initial values of (2) are $1, 0, m-1$, and $e(x)$ increases with $x > 0$; hence f is not universal if $n = 1, 2$, and $f \neq m-2, 2m-3, 3m-4$ if $w < m-2$, $w = m-2$, $w = m-1$ respectively. Hence we let $n = 3$ and $w = m$, $m+1$. Then $a_k \leqq m-1$, and lemmas 1, 2 of R. P. N. hold here, with $w \leqq m+1$. In particular $a_1 = 1$,

$a_2 = 1$ or 2. Since $f \neq 5m - 6$ if $a_1 = 1$ and $a_2 = 2$ we let $a_1 = 1 = a_2$. Hence for $m = 3$ there remains only $(1, 1, 1, 1)$, which is indeed universal finally, since $(1, 1, 1)$ and $(1, 1, 2)$ do not represent 12 and 23 respectively. Again for $m = 4$ with $w = m + 1$ there remains only $(1, 1, 1, 1, 1)$, which is indeed finally universal, since $(1, 1, 1, 2)$ and $(1, 1, 3)$ do not represent 35 and 8 respectively. Otherwise, necessary and sufficient conditions that $f = m - 1, \cdots, m - 1 + w - 1$ are

$$(5) \quad f = (1, 1, a_3, \cdots), \; a_k \leqq w_{k-1} (3 \leqq k \leqq n), \; w = m \geqq 4 \text{ or } w = m + 1 \geqq 6.$$

Next, if f satisfies (5), then necessary and sufficient conditions that $f = 2(m - 1), \cdots, 4(m - 1)$ are

$$(6) \quad \begin{gathered} f = (1, 1, a_3, \cdots), \quad a_3 = 1 \text{ or } 2, \quad a_k \leqq w_{k-1} - 1 \; (4 \leqq k \leqq n), \\ w = m \geqq 4 \quad \text{or} \quad w = m + 1 \geqq 6. \end{gathered}$$

If $w = m = 4$ or 5 in (6), then $f \neq 18$ or 24. Next if f satisfies (6) with $w = m \geqq 6$, then $f \neq 5m - 6$ if $a_3 = 2$ but $f = 4(m - 1), \cdots, 5m - 6$ if $a_3 = 1$. Again, if f satisfies (6) with $w = m \geqq 6$ and $a_3 = 1$, then $f \neq 6m - 9$ if $a_4 = 1$, $a_5 = 3$ but otherwise $f = 5(m - 1), \cdots, 6m - 3$. Finally, since by hypothesis $w \leqq m$ if $m \geqq 6$, we retain of (6) with $w = m + 1$ only those for $m = 5$; these are the last three of (7), and represent $4(m - 1), \cdots, 6m - 3$. Hence we have

THEOREM 1. *Let $w \leqq m$ if $m \geqq 6$ and $w \leqq m + 1$ if $m = 3, 4, 5$. Then $f = 0, \cdots, 6m - 3$ if and only if* (7) *or* (8) *holds:*

$$(7) \quad \begin{matrix} (1,1,1,1), & (1,1,1,1,1), & \\ (1,1,1,1,1,1), & (1,1,1,1,2), & (1,1,2,2), \end{matrix} \qquad w = m = 1,$$

$$(8) \quad \begin{matrix} (1,1,1,a_4,\cdots), \quad a_4 = 1 \text{ or } 2, \\ a_k \leqq w_{k-1} - 1 \; (5 \leqq k \leqq n) \text{ but not } a_4 = 1, \, a_5 = 3. \end{matrix} \qquad w = m \geqq 6,$$

3. *Universal functions of extended polygonal numbers.* Since we use lemmas 10, 12, 13 of R. P. N. in which the variables are positive or zero integers, we here replace (2) by

$$(9) \qquad e(x) = 1 - x + m(x^2 - x)/2, \qquad (x = 0, 1, 2, \cdots).$$

First, let f satisfy (8) with $a_4 = 1$. Then $f = A \geqq 0$ if and only if there are integers a, b, r, each positive or zero, such that $A = r + 4 - b + m(a - b)/2$, where r is represented by f_4, and where there are integers $x, y, z, w \geqq 0$ satisfying $a = x^2 + y^2 + z^2 + w^2$ and $b = x + y + z + w$ whence $e(x) + e(y) + e(z) + e(w) = 4 - b + m(a - b)/2$. As in § 3 of

R. P. N. we find the numbers between 0 and $m-4$ not represented by f_4, and also prove

LEMMA 1. *Let f satisfy* (8) *with $a_4=1$, A be any positive integer, and β be any odd integer $\geqq 5$. Then there are integers a, b, r each $\geqq 0$ such that $A=r+4-b+m(a-b)/2$, $a\equiv b \pmod 2$, $a\not\equiv 0 \pmod 4$, $f_4=r \leqq m-1$, and such that $b=\beta$ or $\beta-2$ if $a_5=1$ but $b=\beta$, $\beta\pm 1$, $\beta-2$, $\beta-3$, $\beta-5$ if $a_5=2$.*

Hence finally f represents A if $A\geqq 44m-48$ for $a_5=1$ and if $A\geqq 296m-80$ for $a_5=2$; the proof is long and follows that of Dickson.* It was verified directly that $f=A$, where $6m-3<A<44m-48$ if $a_5=1$ and $6m-3<A<105m-13$ if $a_5=2$. Hence by theorem 1 we have part of theorem 2.

Next let f satisfy (8) with $a_4=2$. Here we need a, b, r such that $a=x^2+y^2+z^2+2w^2$ and $b=x+y+z+2w$, and prove

LEMMA 2. *Let f satisfy* (8) *with $a_4=2$, A be any positive integer, and β any integer $\geqq 8$. Then there are integers a, b, r each $\geqq 0$ such that $A=r+5-b+m(a-b)/2$, $a\equiv b \pmod 2$, $a\not\equiv 0 \pmod 5$ or $b\not\equiv 0 \pmod 5$, $f_4=r\leqq m-5$, and $b=\beta, \beta-1, \cdots, \beta-8$.*

Hence $f=A\geqq 513m+210$, and we have the second part of

THEOREM 2. *Let f satisfy* (7) *or* (8). *Then f is universal except perhaps for integers A such that $105m-14<A<N$, where $N=296m-80$ for $a_4=1$, $a_5=2$ and $N=513m+210$ if $a_4=2$.*

The preceding proofs hold for (7) with $a_4=1$. For the proof of the universality of the remaining function $(1, 1, 2, 2)$ of theorem 1, we shall use

LEMMA 3. *If a and b are positive odd integers such that $15a\leqq 3b^2+b$ and $b^2\leqq 6a$ then there is a solution in integers x, y, z, w each $\geqq 0$ of $a=x^2+y^2+2z^2+2w^2$ and $b=x+y+2z+2w$.*

For, necessary and sufficient conditions that there be such a solution are that $a\equiv b \pmod 2$ and that there are integers ξ, v, t satisfying

$$(10)\qquad \begin{array}{lll} 6a-b^2=2\xi^2+3v^2+6t^2, & & t\geqq 0, \quad v\geqq 0, \\ 3t+\xi\leqq b, \quad 3v-2\xi\leqq b, & & \xi\equiv b \pmod 3). \end{array}$$

* *American Journal of Mathematics, loc. cit.*, p. 4. By (9) we have

$$e(x)=m[(x-1)^2-(x-1)]/2+(m-1)(x-1),$$

whence we take $c=0$, $t=m-1$, $k=1$; also $D=A-4$, $n=-(m+2)$; and for $a_5=1$ take $d=4$, but $d=8$ for $a_5=2$.

This is evident if we let $x \geqq y$, $z \geqq w$, $u = x + y$, $v = x - y$, $s = z + w$, $t = z - w$, substitute for x from $b = x + y + 2z + 2w$ in $a = x^2 + \cdots + 2w^2$, and write $\xi = b - 3s$. Now if ab is odd and $6a - b^2 \geqq 0$, there are* solutions of (10_1). We choose the sign of ξ so that $\xi \equiv b \pmod 3$; the inequalities in (10) hold if $15a \leqq 3b^2 + b$.

Since our function is $(1, 1, 2, 2)$ we see by (9) that $f = A$ if there are integers a and b, satisfying lemma 3, such that $A = 6 - b + m(a - b)/2$. Let β be any odd integer, and A any positive integer; hence there are integers g, r each $\geqq 0$ such that $A - 6 + \beta = mg + r$, where $0 \leqq r \leqq 4$ (since $w = m + 1$, $m = 5$). Hence if $r = 0, 2, 4$ we take $b = \beta - r$, $a = 2g + b$; but if $r = 1, 3$ then $A - 6 = 5(g - 1) - \beta + r + 5$ and we take $b = \beta - (r + 5)$, $a = 2(g - 1) + \beta - (r + 5)$. Hence always $a \equiv b \equiv 1 \pmod 2$, and $b = \beta, \beta - 2, \cdots, \beta - 8$. Hence, by the preceding method with $d = 10$, we have that $(1, 1, 2, 2)$ is universal if $A \geqq 4792$. Direct verification gives that $f = A < 4792$. This completes the proof of theorem 2.

4. *Necessary conditions for universality in representation by generalized polygonal numbers.* We use (3), and (4) with $f = a_1 g_1 + \cdots + a_n g_n$. By the values of (3) for $|x| = 0, 1, 2$ and the fact that $g(x)$ increases with $|x|$, we see that f is not universal if $n = 1$ or 2, and that $f \neq m - 2$ if $m \geqq 5$ and $w < m - 2$. Hence we let $n \geqq 3$ and $w \geqq m - 2$. But to insure lemmas 1, 2 of R. P. N. we let $w \leqq m - 2$ if $m \geqq 6$; hence $w = m - 2$ if $m \geqq 6$. Similarly, we let $w \leqq m - 1$ if $m = 4$ or 5, and $w \leqq m$ if $m = 3$, as suggested by Dickson's theorem of § 1. But $n \geqq 3$ implies $w \geqq 3$; also if $m = 5$ the functions $(1, 1, 1)$ and $(1, 1, 2)$ do not represent 10 and 23 respectively. Thus if $m = 3, 4, 5$ the only possible universal functions are Dickson's known universal functions. Henceforth, therefore, we let $m \geqq 6$ and $w = m - 2$.

We have lemmas 1, 2 of R. P. N., with $w = m - 2$; that is, $a_1 = 1$, $a_2 = 1$ or 2, and $a_k \leqq w_{k-1} + 1$ $(3 \leqq k \leqq n)$. But $f \neq m$ if $a_1 = 1$ and $a_2 = 2$; again $f \neq m + 1$ if $a_1 = 1 = a_2$ and $a_3 > 2$; also if $a_1 = 1 = a_2 = a_3$, then $f = 2m$ if and only if $a_4 = 1$ or 2; also if $a_1 = 1 = a_2$ and $a_3 = 2$, then $f = 2m + 2$ if and only if $a_4 = 2, 3$, or 4. Otherwise $f = 0, \cdots, 5m + 2$ if and only if

$$
\begin{aligned}
&a_1 = 1 = a_2; \quad w = m - 2 \geqq 4; \\
(11) \qquad &a_3 = 1 \text{ and } a_4 = 1 \text{ or } 2, \quad \textit{or} \quad a_3 = 2 \text{ and } a_4 = 2,\ 3, \text{ or } 4; \\
&n = 4, \quad \textit{or} \quad a_k \leqq w_{k-1} + 1 \ (5 \leqq k \leqq n).
\end{aligned}
$$

If f satisfies (11) with $a_3 = 1 = a_4$, then $f = 5m + 2, \cdots, 34m - 16$.

* W. B. Jones, *Dissertation* (1928), University of Chicago.

In determining those of the remaining functions (11) which represent $5m+2,\cdots,34m-16$ it was necessary to prove general lemmas, similar to lemmas 16 and 17 of R. P. N., and to apply them to the cases $a_3=1$, $a_4=2$, and $a_3=2$ separately.

Let f satisfy (11) with $a_3=1$, $a_4=2$. Then $f=5m+2,\cdots,34m-16$ if $n=4$, or if $n\geqq 5$ and $a_5=2,3,4$. Next let $n\geqq 5$, $a_5=5$. Then $f\neq 5m+5$ if $n=5$, or if $n\geqq 6$ and $a_6=11$; but if $n\geqq 6$ and $a_6=5,\cdots,10$, then $f=5m+2,\cdots,34m-16$. Finally, let $n\geqq 5$, $a_5=6$. If $n=5$, then $f=5m+2,\cdots,34m-16$. If $n>5$, there are four cases: (i) $a_k\neq w_{k-1}$ for every $k\geqq 6$; (ii) $a_k=w_{k-1}$ for some $k\geqq 6$, and $a_k=\cdots=a_n$; (iii) $a_k=w_{k-1}$ for some $k\geqq 6$, and the first coefficient a_K (among $a_{k+1},\cdots,a_n$), which is not equal to a_k, is indeed $>a_k+5$; (iv) $a_k=w_{k-1}$ for at least one $k\geqq 6$, and for every such k there is a coefficient a_K satisfying $a_k<a_K\leqq a_k+5$. Then $f\neq 5m+a_k$ if (ii) or (iii), but $f=5m+2,\cdots,34m-16$ if (i) or (iv). Hence by (11) we have that if $w=m-2\geqq 4$ and $a_1=1=a_2=a_3$, $a_4=2$, then $f=0,\cdots,34m-16$ if and only if $f=(1,1,1,2)$ or

$$
\begin{array}{lll}
 & f=(1,1,1,2,\cdots),\quad a_5=2,3,4; & n=5 \\
 & & \quad or\quad a_k\leqq w_{k-1}+1\ (6\leqq k\leqq n); \\
(12) & f=(1,1,1,2,5,\cdots),\ a_6=5,\cdots,10; & n=6 \\
 & & \quad or\quad a_k\leqq w_{k-1}+1\ (7\leqq k\leqq n); \\
 & f=(1,1,1,2,6,\cdots),\ n=5 & \quad or\quad a_k\leqq w_{k-1}+1\ (6\leqq k\leqq n)
\end{array}
$$

subject to the preceding conditions (i) or (iv).

Next, let f satisfy (11) with $a_3=2=a_4$. Then $f=5m+2,\cdots,34m-16$ if and only if there hold the conditions

(13) $n=4$, or $n>4$ and $a_5=2,3,4$, or 5; or $n>4$, $a_5=6$ or 7, and conditions (i) or (iv) hold.

Finally, if f satisfies (11) with $a_3=2$ and $a_4=4$, then $f=5m+2,\cdots,34m-16$. The same conclusion holds for $a_3=2$, $a_4=3$ if certain conditions, similar to (13), hold; they are not detailed here for the reason noted at the end of § 5.

5. *Certain universal functions of generalized polygonal numbers.* First let f satisfy (11) with $a_3=1=a_4$. If $n=4$ then f is one of Dickson's known universal functions. We prove that f is universal if $n>4$ and

(14) among $a_5,\cdots,a_n$ there is a first coefficient a_J which is not divisible by 4.

Otherwise there is no conclusion. We use

LEMMA 4. *Let f satisfy* (11) *with* $a_3 = 1 = a_4$, *and* (13); *let A be any positive integer and β any odd integer such that* $a_J \leqq \beta \leqq A$. *Then there are integers a, b, r, each* $\geqq 0$, *such that* $A = r + b + m(a - b)/2$, $a \equiv b \pmod 2$, $a \not\equiv 0 \pmod 4$, $f_4 = r \leqq m - 6$, *and b is one of* $\beta - a_J + 1, \beta + a_J + 5$.

Hence we may use the method in § 3 of R. P. N., with $d = 6 + 2a_J$ and $E = m - 6$. Thus we have the first part of theorem 3.

Next let $a_3 = 1$, $a_4 = 2$, and f satisfy (12). We shall prove that f is universal if $n = 4$, or if $n > 4$ and

(15) among $a_5, \cdots, a_n$ there is a first coefficient a_J which is not divisible by 5.

Otherwise there is no conclusion. In order that we may use lemma 13 of R. P. N., we prove

LEMMA 5. *Let f satisfy* (12), *and* (14) *if* $n > 4$; *let A be any positive integer and β any odd integer such that* $a_J \leqq \beta \leqq A$. *Then there are integers a, b, r, each* $\geqq 0$, *such that* $A = r + b + m(a - b)/2$, $a \equiv b \pmod 2$, *a or b is not divisible by 5,* $f_4 = r \leqq m - 7$; *if $n > 4$ then b is one of* $\beta - a_J, \cdots, \beta + 6 + a_J$ *while if $n = 4$ then $r = 0$ and b is one of* $\beta, \cdots, \beta + 13$.

Hence we may use the preceding method, with $d = 15$ and $E = 0$ if $n = 4$ but $d = 8 + 2a_J$ and $E = m - 7$ if $n > 4$. Thus we have the second part of theorem 3.

Next, let f satisfy (13). We shall prove that f is universal if $n = 4$, or if $n > 4$ and

(16) $a_5 = 3, 4, 5,$ or 7; or
$a_5 = 2$ or 6, and among $a_6, \cdots, a_n$ there is a first coefficient a_J which is not divisible by 16.

Otherwise there is no conclusion. We shall use lemma 3 and

LEMMA 6. *If a and b are positive even integers such that* $15a \leqq 3b^2 + b$ *and* $b^2 \leqq 6a$, *and if* (17) *or* (18) *hold, then there are integers x, y, z, w each* $\geqq 0$ *such that* $a = x^2 + y^2 + 2z^2 + 2w^2$ *and* $b = x + y + 2z + 2w$.

The proof of lemma 3 is valid here. But since a and b are even, we require also conditions

(17) $a = 2^{2h-1}A$, $b = 2^iB$, where A and B are each odd and h and i are each $\geqq 1$; and where $h = i$, or $h = i + 1$, or $h + 1 = i$ and $A \not\equiv 1 \pmod 8$, or $h + 1 < i$ and $A \not\equiv 5 \pmod 8$;

(18) $a = 2^{2h}A$, $b = 2^iB$, where A and B are each odd and h and i are each $\geqq 1$; and where $h \leqq i$.

For, if a and b are even, these are the conditions that $6a - b^2$ is not of the form $4^k(8n + 7)$, and hence that (10_1) have solutions. In order that we may use lemmas 3 and 6, we prove

LEMMA 7. *Let f satisfy* (13), *and* (16) *if $n > 4$; let A be any positive integer and β any odd integer such that $5 + a_J \leqq \beta \leqq A$. Then there are integers a, b, r, each $\geqq 0$, such that $A = r + b + m(a - b)/2$, $a \equiv b \pmod 2$, a and b satisfy the hypotheses of lemma 3 or lemma 6, and $f_4 = r \leqq m - 8$; if $n = 4$ then $r = 0$ and b is one of $\beta, \cdots, \beta + 23$; while if $n > 4$ then b is one of $\beta - 6, \cdots, \beta + 5 + a_5$ if $a_5 = 3, 5, 7$, but if $a_5 = 4$ then b is one of $\beta - 9, \cdots, \beta + 6$, while if $a_5 = 2, 6$ then b is one of $\beta - 5 - a_J, \cdots, \beta - 1 + a_5 + a_J$.*

Hence we may use the preceding method, with $E = 0$ and $d = 25$ if $n = 4$, but with $E = m - 8$ if $n > 4$ where $d = a_5 + 13$ if $a_5 = 3, 5, 7$, and $d = 17$ if $a_5 = 4$, and $d = 6 + a_5 + 2a_J$ if $a_5 = 2, 6$. This proves the last part of

THEOREM 3. *Let f satisfy* (11) *with $a_3 = a_4 = 1$, and also* (14) *if $n > 4$; or let f satisfy* (12), *and also* (15) *if $n > 4$; or let f satisfy* (13), *and also* (16) *if $n > 4$. Then f is universal except perhaps for integers A such that $34m - 16 < A < N$, where*

$$N = 2m(14a_J^2 + 63a_J + 71) + 28(a_J + 1) \quad \text{if } a_3 = a_4 = 1 \quad \text{and } n > 4,$$
$$N = 36m(4 + a_J)^2 + 36(a_J + 2) \quad \text{if } a_3 = 1, a_4 = 2 \quad \text{and } n > 4,$$
$$N = m(11d^2 - 36d + 39) + 22d - 119 \quad \text{if } a_3 = 2 = a_4 \quad \text{and } n > 4;$$

if $n = 4$, f is universal except perhaps for integers A such that $1000 < A < N$ *where $N = 11,711$ if* $(1, 1, 1, 2)$ *and $N = 48,439$ if* $(1, 1, 2, 2)$.

Finally, we are unable to complete the general proof if $a_3 = 2$ when $a_4 = 3, 4$, since it has been impossible as yet to prove for these cases lemmas analogous to lemmas 3 and 6.

NORTHWESTERN UNIVERSITY.

ON THE POSSIBLE FORMS OF DISCRIMINANTS OF ALGEBRAIC FIELDS. II.

By William R. Thompson.

In a previous communication * a complete solution has been given to the problem of finding the powers of a given rational prime which may divide exactly the discriminant of an algebraic field of n-th degree. The foundation of the proof lay in a report by Ore; † and the existence of a later report ‡ by the same author, giving similar data for relative fields, suggested the possibility of ascertaining in similar manner what powers of a given prime-ideal of order, t, may divide exactly the relative discriminant of a relative field of n-th degree.

It is the object of the present communication to give the solution to this problem. The similarity of the demonstrations required is so marked that the previous form may be used extensively to indicate processes in the proof of this more general case, and the statement of the results may be made in almost as concise a form as in the case previously treated.

For convenience let us refer to this previous work § simply as *Part I* and the present as *Part II*; and let $\phi_{(\omega)}$ be an algebraic field (called the *fundamental field*), p a rational prime greater than 1, and $\mathfrak{P}$ a prime-ideal divisor of p of order t in $\phi_{(\omega)}$.

Now, let $K_{(\theta)}$ be a relative field of n-th degree; i. e., let θ be a root of an irreducible equation

$$(1) \qquad g_{(x)} = \sum_{i=0}^{n} a_i x^i = 0$$

* W. R. Thompson, *American Journal of Mathematics*, Vol. 53 (1931), pp. 81-90. In what follows, this will be referred to as Paper I.

† O. Ore, *Mathematische Annalen*, Vol. 96 (1926), pp. 313-352. In what follows, this will be referred to as Paper II.

‡ O. Ore, *Mathematische Annalen*, Vol. 97 (1927), pp. 569-598. A comparison between this article of Ore and the present communication may be facilitated by the translation,

$$\begin{array}{ll} \textit{Present Notation,} & \phi_{(\omega)},\ K_{(\theta)},\ \mathfrak{P},\ t,\ P,\ n,\ d,\ e,\ f,\ \rho; \\ \textit{Ore's Notation,} & k_{(\theta)},\ K_{(\Omega)},\ P,\ e,\ \mathfrak{P},\ m,\ D_k,\ E,\ F,\ R; \end{array}$$

the present notation being in accord with that of the previously mentioned Papers I and II, Ore's later notation being introduced in dealing simultaneously with *ordinary* and *relative supplemental numbers* in developing the *Verzweigungstheorie* upon which the present work is based. In what follows, this will be referred to as Paper III.

§ Paper I.

where the coefficients (a_i) are integers of the field $\phi_{(\omega)}$ and $a_n = 1$. Let d be the *relative* discriminant of $K_{(\theta)}$ and let $\varepsilon \geq 0$ be defined as a rational integer such that d is exactly divisible by $\mathfrak{P}^{\varepsilon}$. Then the maximal value of ε which is attainable for such relative fields has been given by Ore.*

Accordingly, if we let M be this maximal value and let n be given *p-adically* by

$$n = \sum_{a=0}^{q} b_a p^a, \quad \text{where} \quad 0 \leq b_a < p \tag{2}$$

and b_a is a rational integer; and J is the aggregate number of these coefficients (b_a) which are different from zero; and let $N_{(n,p,t)}$ be defined by

$$N_{(n,p,t)} = n - J + t \cdot \sum_{a=0}^{q} a b_a p^a; \tag{3}$$

then Ore has shown that $M = N_{(n,p,t)}$. Furthermore, it is obvious from the definition contained in (3) and (4) of *Part I* that if $t = 1$ then $N_{(n,p,t)} = N_{(n,p)}$.

Now, in the relative field, $K_{(\theta)}$, let the prime-ideal decomposition of the ideal, $\mathfrak{P}$, be given by

$$\mathfrak{P} = \prod_{i=1}^{r} P_i^{e_i}, \qquad N_{(P_i)} = \mathfrak{P}^{f_i} \tag{4}$$

where $N_{(P_i)}$ is the *relative* norm of P_i with respect to $\phi_{(\omega)}$. Then e_i and f_i are called, respectively, the *relative order* and *relative degree* of the prime-ideal P_i; and (as is well known) there exists the relation

$$n = \sum_{i=1}^{r} e_i f_i \quad \text{and} \quad e_i > 0 < f_i. \tag{5}$$

Now, let ρ_i be the *relative supplemental number of* P_i as defined by Ore;* then we may state for relative fields what we shall call *Ore's Third Theorem* which is contained in the paper previously mentioned.*

ORE'S THIRD THEOREM. *For each prime-ideal, P_i, there exists a rational integer, $\rho_i \geq 0$, such that if $S_i \geq 0$ be a rational integer such that e_i is exactly divisible by p^{S_i} then ρ_i is determined as follows:*

$$\begin{aligned} &\textit{if } S_i = 0, \textit{ then } && \rho_i = 0; \\ &\textit{and if } S_i \neq 0, \textit{ then } && 1 \leq \rho_i \leq t e_i S_i, \end{aligned} \tag{6}$$

and in this latter alternative ρ_i is restricted by the condition that if there

* Paper III.

exists a positive rational integer, ν_i, *such that* ρ_i *is exactly divisible by* p^{ν_i}, *then* ν_i *shall not exceed* $\rho_i/t \cdot e_i$; *and then* ε *is given by*

$$(7) \qquad \varepsilon = \sum_{i=1}^{r} f_i(e_i - 1 + \rho_i);$$

and for any set of rational integers, designated by

$$(8) \qquad \left\{ \begin{matrix} e_1, \cdots, e_r \\ f_1, \cdots, f_r \\ \rho_1, \cdots, \rho_r \end{matrix} \right\}$$

and satisfying the conditions of (5) *and* (6) *there exists an algebraic field of* n*-th degree relative to* $\phi_{(\omega)}$ *such that its relative discriminant is exactly divisible by* $\mathfrak{P}$ *where* ε *has the value given in* (7). *Furthermore, the prime-ideal decomposition of* $\mathfrak{P}$ *is given by* (4) *in the relative field corresponding to* (8), *and the maximal value of* ε *attainable for any such field of* n*-th degree relative to* $\phi_{(\omega)}$ *wherein* $\mathfrak{P}$ *has the order,* t, *is* $N_{(n,p,t)}$.

Obviously, the conclusions of this theorem apply equally to any other ideal divisor of p which is prime in $\phi_{(\omega)}$ and has the same order t.

Now, let $\mathcal{E}^{(n)}_{p,t}$ be defined as the class exactly containing the attainable values of ε for a relative field of n-th degree as above. Obviously, as an algebraic field can be regarded in any case as a relative field with respect to the rational field, it follows that *Part I* had to deal with merely a special case of the present problem which is essentially that of evaluating the components of $\mathcal{E}^{(n)}_{p,t}$; i. e., $\mathcal{E}^{(n)}_{p,1} = \mathcal{E}_p^{(n)}$, defined in *Part I*. Furthermore, it is obvious that the components of the class, $\mathcal{E}^{(n)}_{p,t}$, depend entirely upon n, p and t.

In the same manner as for the special case,* $t = 1$, we may prove the

THEOREM 9. $\mathcal{E}^{(n)}_{p,t}$ *is a set of a finite number of rational integers including* 0 *and* $N_{(n,p,t)}$ *as least and greatest component, respectively.*
and the

THEOREM 10. *For* $p > n$, $\mathcal{E}^{(n)}_{p,t} = 0, \cdots, N_{(n,p,t)}$.

Now, let a set of numbers arranged as in (8), satisfying the prescribed conditions, be called a *relative critical matrix.* Then, by the same devices as employed * for critical matrices in *Part I,* we prove the

THEOREM 11. *If* n' *and* n'' *are two positive rational integers such that* $n' + n'' = n$, *then*

* Paper I.

$$\mathcal{E}^{(n)}_{p,t} \quad \textit{includes} \quad \mathcal{E}^{(n')}_{p,t} + \mathcal{E}^{(n'')}_{p,t}$$
$$\textit{and} \quad \mathcal{E}^{(n)}_{p,t} \quad \textit{includes} \quad \mathcal{E}^{(n')}_{p,t}.$$

Here, of course, the sum of two classes has the special meaning given in (16) of *Part I.*

Now, let the sets, $A^{(n)}_{p,t}$ and $H^{(n)}_{p,t}$, be called the *acquisition* and the *heritage*, respectively, of $\mathcal{E}^{(n)}_{p,t}$ and be defined as follows:

$A^{(n)}_{p,t}$ shall contain exactly those components of $\mathcal{E}^{(n)}_{p,t}$ which are not components of the *sum* of any two classes, $\mathcal{E}^{(n')}_{p,t}$ and $\mathcal{E}^{(n'')}_{p,t}$, where n' and n'' are *positive* rational integers such that $n = n' + n''$; and $H^{(n)}_{p,t}$ shall contain exactly all other components of $\mathcal{E}^{(n)}_{p,t}$.

Then, in the same manner as for the special case which has been treated in *Part I* (reference to which may be made in the present notation by merely specifying the case, $t = 1$, for which the treatment in *Part I* applies exactly), we prove the

THEOREM 12. *If ε is a component of $A^{(n)}_{p,t}$, it corresponds to a relative critical matrix of the form*

$$\left\{ \begin{matrix} n \\ 1 \\ \rho \end{matrix} \right\} \quad \textit{and} \quad \varepsilon = n - 1 + \rho$$

where ρ is a rational integer defined by the relations:

If $S \geq 0$ is a rational integer such that n is exactly divisible by p^S, then

$$\textit{if } S = 0, \qquad \rho = 0,$$
$$\textit{and if } S \neq 0, \qquad 1 \leq \rho \leq tnS,$$

and in this latter alternative ρ is restricted by the condition that if there exists a positive rational integer, v, such that ρ is exactly divisible by p^v, then $v \leq \rho/t \cdot n$.

Accordingly, by the definition of $A^{(n)}_{p,t}$ and *Ore's Third Theorem* we have

THEOREM 13. *That ε be a component of $A^{(n)}_{p,t}$ it is necessary and sufficient that the conditions of Theorem 12 be satisfied and that ε shall not be a component of any sum class of the type, $\mathcal{E}^{(n')}_{p,t} + \mathcal{E}^{(n'')}_{p,t}$, where n' and n'' are positive rational integers whose sum is n.*

Now, consider the case, $n = p$. Then, obviously, by *Theorem* 10 and the definition of the heritage of $\mathcal{E}^{(n)}_{p,t}$ we have (just as for $t = 1$)

$$H^{(p)}_{p,t} = 0, \cdot \cdot \cdot, p - 2; \tag{9}$$

and for ε in the acquisition, by *Theorem* 12, we have for $n = p$, then $S = 1$, and $1 \leq \rho \leq tn$, whence if $\rho \equiv 0 \pmod{p}$ then $\rho = tn$, whence we have [as $N_{(p,p,t)} = p - 1 + tp$]

$$(10) \qquad A^{(p)}_{p,t} = p, \cdots, N_{(p,p,t)}$$

except all numbers of the form $gp - 1$, where g is a positive rational integer less than $t + 1$; and (9) and (10) give the

THEOREM 14. $\mathcal{E}^{(p)}_{p,t} = 0, \cdots, N_{(p,p,t)}$ *except all numbers of the form* $gp - 1$, *where* $1 \leq g \leq t$, *and* g *is a rational integer.*

We note that in particular for $p = 2$ we have

$$(11) \qquad \mathcal{E}^{(2)}_{2,t} = 0, \cdots, 2t + 1, \text{ except all odd numbers less than } 2t.$$

As before,* we define an *exceptional number* relative to $\mathcal{E}^{(n)}_{p,t}$ as any rational integer, η, not in $\mathcal{E}^{(n)}_{p,t}$ and such that $0 \leq \eta \leq N_{(n,p,t)}$. Then by *Theorem* 9, obviously, $0 < \eta < N_{(n,p,t)}$.

By complete induction we show that the odd numbers less than $2t$ are exceptions relative to $\mathcal{E}^{(n)}_{2,t}$ for every n. Obviously, this is true for n less than 3; and if it is true for any $n < m$, where m is a rational integer, then, obviously, $H^{(m)}_{2,t}$ contains no odd number less than $2t$; and by Theorems 12 and 13 if ε is an odd number less than $2t$ and in $A^{(m)}_{2,t}$ then m must be even and, as $\varepsilon = m - 1 + \rho$, then $\rho < 2t \leqq tm$, whence $\rho/t \cdot m < 1$ whence ρ is odd, which is impossible. Accordingly, we have proved the

THEOREM 15. *The odd numbers less than* $2t$ *are never components of* $\mathcal{E}^{(n)}_{2,t}$ *for any* n.

Let η be called a *universal exception* relative to the prime, p, and the order number, t, if and only if η is not a component of $\mathcal{E}^{(n)}_{p,t}$ for any n. Then we may restate

THEOREM 15. *The odd numbers less than* $2t$ *are universal exceptions relative to* 2 *and the order* t.

Now, if η is an exceptional number relative to $\mathcal{E}^{(n)}_{p,t}$ but $\eta - 2$, $\eta - 1$, $\eta + 1$ and $\eta + 2$ *are* components of $\mathcal{E}^{(n)}_{p,t}$; then and only then let η be called a *regular exception* relative to $\mathcal{E}^{(n)}_{p,t}$, other exceptional numbers to be called *irregular*. This is merely a generalization of the definition employed in *Part I*. Thus we may state the

* Paper I.

THEOREM 16. *If n' and n'' are two positive rational integers and $n' + n'' = n$, and the sets, $\mathcal{E}^{(n')}_{p,t}$ and $\mathcal{E}^{(n'')}_{p,t}$ have no irregular exceptions unless $p = 2$ and in that case the only irregular exceptions are the universal exceptions relative to 2 and the order t; then, if $n' > 1 < n''$, then*

$$\mathcal{E}^{(n')}_{p,t} + \mathcal{E}^{(n'')}_{p,t} = 0, \cdots, (N_{(n',p,t)} + N_{(n'',p,t)})$$

except the universal exceptions if $p = 2$, and if either n' or $n'' = 1$, then

$$\mathcal{E}^{(n')}_{p,t} + \mathcal{E}^{(n'')}_{p,t} = \mathcal{E}^{(n-1)}_{p,t}.$$

In view of the definition of *class-summation* employed [as given in (16) of *Part I*] and *Theorem* 9, the proof is obvious (as the only component of $\mathcal{E}^{(1)}_{p,t}$ is zero). We are now ready to state and prove by the method of complete induction the main theorem,

THEOREM 17. *If $p > 2$ and α is a positive rational integer; then if $n = p^{\alpha}$, $\mathcal{E}^{(n)}_{p,t} = 0, \cdots, N_{(n,p,t)}$ except all numbers of the form $(\alpha t p^{\alpha} - 1 - g p^{\alpha})$ where g is a rational integer and $0 \leq g < t$, if $n = p^{\alpha} + 1$, $\mathcal{E}^{(n)}_{p,t} =$ Union $[(p-1)$ and $\mathcal{E}^{(n-1)}_{p,t}]$, and in every other case $\mathcal{E}^{(n)}_{p,t} = 0, \cdots, N_{(n,p,t)}$; and if $p = 2$, then $\mathcal{E}^{(n)}_{p,t}$ is formally as given above (for the case $p > 2$) except that the odd numbers less than $2t$ are never components of $\mathcal{E}^{(n)}_{2,t}$.*

By a consideration of the same cases and in the same order as in the proof of *Theorem* 8 (in *Part I*) we may demonstrate readily that if k is a positive rational integer such that *Theorem* 17 is verified for every $n \leq p^k$ then it may be verified for every $n \leq p^{k+1}$. By *Theorems* 10 and 14 there exists at least one value for k, namely 1; hence we may assume k to be such a number. Obviously, it suffices to prove (in addition to the above) that *Theorem* 17 may be verified for $p^k + 1 \leq n \leq p^{k+1}$ in order to establish the theorem by the method of complete induction. Obviously, we may refer to *Theorem* 17 for the components of $\mathcal{E}^{(\mu)}_{p,t}$ in any case for which the theorem has been verified in the course of the proof, and *à fortiori* for $\mu \leq p^k$, where μ is a positive rational integer. Obviously, then (for such values of μ)

(12) $\mathcal{E}^{(\mu)}_{p,t}$ is without *irregular exception* unless $p = 2$ and then the only irregular exceptions are the *universal exceptions* relative to 2 and the order t, namely, the odd numbers less than $2t$.

Consider the case, $n = p^k + 1$. Then by (2) and (3) we have

$$N_{(n,p,t)} = N_{(p^k,p,t)} = p^k - 1 + tkp^k.$$

Obviously, if $p = 2$ and $k = 1$, then

$$\mathcal{E}^{(n)}_{p,t} = \mathcal{E}^{(p^k)}_{p,t} = \mathcal{E}^{(2)}_{2,t};$$

as the exceptions involved in this case are *universal.* However, if $p > 2$ or $k > 1$, then the least regular exception (as given by *Theorem* 17) for $\mathcal{E}^{(p^k)}_{p,t}$ is $tkp^k - 1 - (t-1)p^k$, obviously, which we shall represent by L. Then, if M is the greatest component of a class $\bar{H}^{(n)}_{p,t}$ defined in the same manner as $H^{(n)}_{p,t}$ except that n' and n'' are restricted to values greater than 1; we have by relation (12) and *Theorem* 16 (the value of $N_{(\mu,p,t)}$ for any positive rational integer, μ, being given by relations (2) and (3)) that for the present case ($n = p^k + 1$, where $p > 2$ or $k > 1$)

$$(13) \quad M = n - 2 + t(k-1)\cdot p \cdot p^{k-1} - \lambda_k, \text{ where } \lambda_k \begin{Bmatrix} = 0 \text{ if } k = 1 \\ = 1 \text{ if } k > 1 \end{Bmatrix};$$

and therefore $L - M = \lambda_k$.

Obviously, then $(p-1)$ is a component of $\mathcal{E}^{(p+1)}_{p,t}$ if $p > 2$ (if $p = 2$ then $p - 1$ is always an exceptional number); and by *Theorems* 11 and 13 (as the acquisition is void) obviously *Theorem* 17 is verified for the case, $n = p^k + 1$.

Now, consider the case, $n = b \cdot p^k$ where b is a rational integer and $1 < b < p$. Obviously, then $p > 2$. Then by the definition of k, (12) and *Theorem* 16 if we set $n' = (b-1)p^k$ and $n'' = p^k$ then if $\mathcal{E}^{(n')}_{p,t}$ is as given in *Theorem* 17 we have, obviously,

$$(14) \qquad H^{(n)}_{p,t} \textit{ includes } 0, \cdots, N_{(n',p,t)} + N_{(n'',p,t)},$$

but by (2) and (3) the last-mentioned component equals $N_{(n,p,t)} - 1$ in this case, whence *Theorem* 9 gives (for the conditions stated)

$$(15) \qquad \mathcal{E}^{(n)}_{p,t} = 0, \cdots, N_{(n,p,t)}.$$

Obviously, the *proviso* that $\mathcal{E}^{(n')}_{p,t}$ be as given in *Theorem* 17 is satisfied in the case, $b = 2$, by the definition of k; whence by complete induction *Theorem* 17 is verified for the case, $n = b \cdot p^k$ for any positive rational integer, $b < p$.

Now, consider the case $p^k + 1 < n < p^{k+1}$ where $n \not\equiv 0 \pmod{p^k}$. Then n may be expressed in p-adic form by (2) where $q = k$ by

$$(16) \qquad n = \sum_{a=0}^{k} b_a p^a, \text{ where } 0 \leq b_a < p$$

and b_a is a rational integer. Now, let $n' = b_k \cdot p^k$ and $n'' = n - n'$. Obviously, then $n'' < p^k$, and $n' > 1 < n''$ whence $\mathcal{E}^{(n')}_{p,t}$ and $\mathcal{E}^{(n'')}_{p,t}$ are as given in *Theorem* 17; and, furthermore, $N_{(n',p,t)} + N_{(n'',p,t)} = N_{(n,p,t)}$; whence (12) and *Theorem* 16 suffice to prove *Theorem* 17 is verified in the given case, and by previous proof in combination with this the theorem is verified for

every $n < p^{k+1}$. Thus it only remains to establish the theorem for the case, $n = p^{k+1}$.

Accordingly, consider the case, $n = p^{k'}$ where $k' = k + 1$. Let M be the maximal value in this case attained for the sum, $N_{(n',p,t)} + N_{(n'',p,t)}$, for any positive rational integers, n' and n'' such that $n' + n'' = n$. Obviously, this value is attained when $n' = (p - b)p^k$ and $n'' = bp^k$ where b is any positive rational integer less than p. Thus $M = n - 2 + tkp^{k+1} = k'tp^{k'} - 2 - (t - 1)p^{k'}$; and (12) and *Theorem* 16 give

$$(17) \qquad H^{(p^{k'})}_{p,t} = 0, \cdots, k'tp^{k'} - 2 - (t-1)p^{k'} \text{ except, if } p = 2, \text{ the odd numbers less than } 2t.$$

Therefore, by *Theorems* 15, 12 and 13 as $N_{(p^{k'},p,t)} = k'tp^{k'} - 1 + p^{k'}$ we have

$$(18) \qquad A^{(p^{k'})}_{p,t} = k'tp^{k'} - (t-1)p^{k'}, \cdots, N_{(p^{k'},p,t)}$$

except all numbers of the form $k'tp^{k'} - 1 - gp^{k'}$ *where* g *is a rational integer and* $0 \leq g < t - 1$.

Accordingly, $\mathcal{E}^{(p^{k+1})}_{p,t}$ is as given in *Theorem* 17; whence by complete induction the theorem is completely verified. It may be observed readily that all possible cases are covered, and that when the fundamental field $(\phi_{(\omega)})$ is rational *Theorem* 17 reduces to *Theorem* 8 of *Part I*.

YALE UNIVERSITY.

THE VOLTAGE INDUCED IN A TRANSMISSION LINE BY A LIGHTNING DISCHARGE.

By F. H. Murray.

In theoretical investigations of the effects of a lightning discharge on a neighboring transmission line, it is often sufficient to consider only the redistribution of the charge bound on the line by the charged cloud, after the cloud discharge has commenced. It is the purpose of this paper to develop a formal solution of the electrodynamical problem for a non-terminated line, with the assumption of a point cloud and a perfectly conducting earth, the cloud discharge being assumed to be that of a condenser; the usual transmission line equations are extended to include the varying impressed electric field which is present after the cloud discharge has begun, and these are solved by the method of Riemann. From this solution is obtained an approximate solution for the horizontal rectangular cloud with constant surface density.

For an ideal line with no resistance or conductance, the approximate form of the voltage wave crest at a distance is obtained; for a point cloud the voltage jumps to a maximum and falls off exponentially; for a horizontal cloud the voltage rises continuously during an interval $\Delta t = l/c$, l being the cloud length parallel to the line, afterwards falling off exponentially.*

1. *The transmission line equations.* Let the point cloud have the charge Q, and coördinates $(0, \bar{y}, \bar{z})$, its image has the charge $-Q$ and coördinates $(0, \bar{y}, -\bar{z})$; if r_1, r'_1 are the distances of an arbitrary point from the cloud and its image, respectively, the static electric field has the potential

$$\Phi_0 = (Q/4\pi)(1/r_1 - 1/r'_1). \tag{1.1}$$

The horizontal conductor is taken to be the line $y = 0$, $z = h$; this line and its image $(y = 0,\ z = -h)$ form the transmission line of distributed constants L, R, C, G. From symmetry, the charges at image points of the line are of opposite sign, the currents are equal in magnitude but opposite in direction. If $E_x{}^0$, $E_z{}^0$ are the components of the impressed field, and

$$E_x = 2E_x{}^0, \quad E_z = \int_{-h}^{h} E_z{}^0 dz \tag{1.2}$$

* See "Traveling Waves Due to Lightning," L. V. Bewley, *Transactions of the American Institute of Electrical Engineers*, July, 1929, p. 1050, for an approximate discussion of terminated lines; for a general discussion of the lightning discharge, F. W. Peek, Jr., "Dielectric Phenomena in High-Voltage Engineering," 1929.

the extended transmission line equations are easily seen to be *

$$(1.3)\qquad \begin{aligned}(L\partial/\partial t + R)I &= -\partial V/\partial x + E_x,\\ C\partial V/\partial t + G(V - E_z) &= -\partial I/\partial x.\end{aligned}$$

The initial static distribution is defined by

$$I = 0,\ V = E_z,\ -\partial V/\partial x + E_x = 0,$$

the last equation resulting from the preceding one and the existence of the potential Φ_0 for the impressed field. Eliminating I and V, respectively,

$$(1.4)\qquad \begin{aligned}(L\partial/\partial t + R)(C\partial/\partial t + G)V - \partial^2 V/\partial x^2 &= -\partial E_x/\partial x + G(L\partial/\partial t + R)E_z,\\ (L\partial/\partial t + R)(C\partial/\partial t + G)I - \partial^2 I/\partial x^2 &= -G\partial E_z/\partial x + (C\partial/\partial t + G)E_x.\end{aligned}$$

Each of these is of the form

$$(1.5)\qquad [LC\partial^2/\partial t^2 + (LG + RC)\partial/\partial t + RG]u - \partial^2 u/\partial x^2 = f(x, t).$$

If $\qquad v = (LC)^{-1/2},\ \alpha = R/2L,\ \beta = G/2C,\ \rho = \alpha + \beta,\ \sigma = \alpha - \beta \neq 0,$

$$\tau = \sigma t,\ y = \sigma x/v,\ u = e^{-\rho t}W,$$

the resulting equation

$$(1.6)\qquad \partial^2 W/\partial y^2 - \partial^2 W/\partial \tau^2 + W = -(v/\sigma)^2 f e^{\rho\tau/\sigma} = g(y, \tau)$$

can be integrated by the method of Riemann.† If

$$f(y) = W\,|_{\tau=0},\ F(y) = \partial W/\partial \tau\,|_{\tau=0},$$

$$u = [(\tau - \tau_1)^2 - (y - y_1)^2]^{1/2},\ \bar{v} = I_0(u) = \sum_{n=0}^{\infty} (u/2)^{2n}/(n!)^2$$

the solution is,

$$\begin{aligned}2W(y_1, \tau_1) = f(y_1 - \tau_1) + f(y_1 + \tau_1)&\\ + \int_{y_1-\tau_1}^{y_1+\tau_1} \bar{v}F\,dy + \tau_1 \int_{y_1-\tau_1}^{y_1+\tau_1} (1/u)(d\bar{v}/du)f\,dy&\\ - \int_0^{\tau_1} d\tau \int_{y_1+\tau-\tau_1}^{y_1+\tau_1-\tau} \bar{v}g(y, \tau)\,dy.*&\end{aligned}$$

In the present problem let the cloud discharge begin at the time $t = 0$; both I and V are constant for a small time interval, hence

$$\partial U/\partial t\,|_{t=0} = 0 = (-\rho W + \partial W/\partial t)\,|_{t=0}$$

and

---

* J. R. Carson and R. S. Hoyt, "Propagation of Periodic Currents Over a System of Parallel Wires," Appendix I, *The Bell System Technical Journal*, Vol. 6 (July, 1927), pp. 495-545.

† Riemann-Weber, *Die Partiellen Differentialgleichungen der Mathematischen Physik*, Vol. II, 1901, p. 310.

$$
\begin{aligned}
2U(x_1, t_1)e^{\rho t_1} &= U_0(x_1 - vt_1) + U_0(x_1 + vt_1) + v^{-1}\int_{x_1-vt_1}^{x_1+vt_1} \bar{v}\rho U_0 dx \\
&\quad + \sigma^2 t_1 v^{-1}\int_{x_1-vt_1}^{x_1+vt_1} u^{-1}(d\bar{v}/du)U_0 dx + v\int_0^{t_1} e^{\rho t}dt \int_{x_1-v(t_1-t)}^{x_1+v(t_1-t)} \bar{v}f(x,t)\,dx.
\end{aligned} \tag{1.7}
$$

2. *Discussion of the solution.* The field impressed on the line may be expressed in terms of scalar and vector potentials:

$$
\begin{aligned}
&\boldsymbol{E}^0 = -\operatorname{grad}\Phi - (1/c)\partial\boldsymbol{A}/\partial t, \quad \boldsymbol{H}^0 = (1/c)\operatorname{curl}\boldsymbol{A}. \\
&4\pi\Phi = r_1^{-1}Q_{t-r_1/c} - r'^{-1}_1 Q_{t-r'_1/c}, \\
&4\pi c\boldsymbol{A} = \int_{-\bar{z}}^{\bar{z}} r^{-1}\boldsymbol{I}_{t-r/c}dz.
\end{aligned}
$$

For a condenser discharge,

$$
\begin{aligned}
&Q = Q_0,\ t < 0,\ Q(t) = e^{-\gamma t}Q_0,\ t > 0, \\
&I = dQ/dt = -\gamma Q(t),\ t > 0,\ I = 0,\ t < 0,
\end{aligned}
$$

consequently,

$$
\begin{aligned}
&4\pi\Phi = Q_0[e^{-\gamma(t-r_1/c)}/r_1 - e^{-\gamma(t-r'_1/c)}/r'_1] \\
&4\pi c A_z = -\gamma Q_0 e^{-\gamma t}\int_{-\bar{z}}^{\bar{z}} \begin{bmatrix} 0,\ r > ct \\ e^{\gamma r/c},\ r < ct \end{bmatrix} dz'/r, \quad A_x = A_y = 0.
\end{aligned}
$$

In the following analysis let $G = 0$, and let $w = x_1 + v(t_1 - t)$, $\bar{w} = x_1 - v(t_1 - t)$; from (1.4), (1.7),

$$
2I(x_1, t_1) = Cve^{-\rho t_1}\int_0^{t_1} e^{\rho t}dt \int_{\bar{w}}^{w} \bar{v}(\partial E_x/\partial t)\,dx,
$$

hence

$$
I(x_1, t_1) = -Cve^{-\rho t_1}\int_0^{t_1} e^{\rho t}dt \int_{\bar{w}}^{w} \bar{v}(\partial^2\Phi/\partial x\partial t)\,dx.
$$

Integrating by parts,

$$
\begin{aligned}
I(x_1, t_1)e^{\rho t_1}/Cv &= -\int_0^{t_1} e^{\rho t}\,[\partial\Phi/\partial t\,|_{x=w} - \partial\Phi/\partial t\,|_{x=\bar{w}}]\,dt \\
&\quad + \int_{x_1-vt_1}^{x_1+vt_1} [\Phi e^{\rho t}\partial\bar{v}/\partial x]_{t=0}^{t=t_1-|x-x_1|/v} dx - \iint_{|x-x_1|<v(t_1-t)} \Phi e^{\rho t}(\rho\partial\bar{v}/\partial x + \partial^2\bar{v}/\partial t\partial x)\,dxdt
\end{aligned} \tag{2.1}
$$

and

$$
\begin{aligned}
V(x_1, t_1)e^{\rho t_1} &= -\Phi_0(x_1 - vt_1) - \Phi_0(x_1 + vt_1) - (\rho/v)\int_{x_1-vt_1}^{x_1+vt_1} \bar{v}\Phi_0(x)\,dx \\
&\quad -(\sigma^2 t_1/v)\int_{x_1-vt_1}^{x_1+vt_1} u^{-1}(d\bar{v}/du)\Phi_0(x)\,dx \\
&\quad + v\int_0^{t_1} e^{\rho t}[\Phi_x(w, t) - \Phi_x(\bar{w}, t)]\,dt
\end{aligned} \tag{2.2}
$$

$$+\sigma^2/2\int_0^{t_1} e^{\rho t}(t_1-t)[\Phi(w,t)+\Phi(\bar{w},t)]\,dt$$

$$+v\int_0^{t_1} e^{\rho t}dt\int_{\bar{w}}^{w}\Phi\partial^2\bar{v}/\partial x^2\,dx.$$

with the aid of identities

$$\bar{v}=1,\quad \partial\bar{v}/\partial x=\mp(\sigma^2/2v)(t_1-t),\quad x=x_1\pm v(t_1-t).$$

Let ϕ, ϕ' be the retarded potentials of the cloud and its image, respectively; it follows from (2.1), (2.2) that I and V are of the form $I=I_1-I_2$, $V=V_1-V_2$, I_1, V_1 being represented in terms of ϕ. To simplify the analysis let $R=0$, from which $\rho=\sigma=0$.

3. *Formulae for the voltage.* If $\rho=\sigma=0$, (2.2) gives for V_1,

$$V_1(x_1,t_1)=-[\phi_0(x_1-vt_1)+\phi_0(x_1+vt_1)]$$
$$+v\int_0^{t_1}(\partial/\partial x_1)[\phi(w,t)-\phi(\bar{w},t)]\,dt.$$

Now $\phi(w,t)=\phi_0(w)$ for $0<t<\bar{t}_1$, if $\bar{t}_1$ is defined by the equation

$$r=c\bar{t}_1,\quad x=w_1=x_1+v(t_1-\bar{t}_1),\quad r^2=(\bar{z}-h)^2+\bar{y}^2+x^2$$

and

$$\phi(\bar{w},t)=\phi_0(\bar{w}),\quad 0<t<\bar{t}_2,\quad c\bar{t}_2=r,\quad x=x_1-v(t_1-\bar{t}_2)=\bar{w}_1$$

$$v\int_0^{\bar{t}_1}(\partial/\partial x_1)\phi_0(w)\,dt=-\int_0^{\bar{t}_1}(\partial/\partial t)\phi_0(w)\,dt=\phi_0(x_1+vt_1)-\phi_0(w_1)$$

$$-v\int_0^{\bar{t}_2}(\partial/\partial x_1)\phi_0(\bar{w})\,dt=\phi_0(x_1-vt_1)-\phi_0(\bar{w}_1).$$

Hence

$$V_1(x_1,t_1)=-[\phi_0(w_1)+\phi_0(\bar{w}_1)]+v\int_{\bar{t}_1}^{t_1}\phi_{x_1}(w,t)\,dt-v\int_{\bar{t}_2}^{t_1}\phi_{x_1}(\bar{w},t)\,dt.$$

To evaluate the first integral,

$$\int_{\bar{t}_1}^{t_1}\phi(w,t)\,dt=(Q_0/4\pi)\int_{\bar{t}_1}^{t_1}e^{-\gamma(t-r/c)}dt/r=(Q_0/4\pi)I_1$$

$$I_1=-v^{-1}e^{-\gamma(t_1+x_1/v)}\int_{w_1}^{w_2}e^{\gamma(w/v+r/c)}dw/r\qquad w_2=x_1.$$

Let $d^2=(\bar{z}-h)^2+\bar{y}^2$, $r^2=d^2+w^2$,

$$du=r+w,\quad r=(d/2)(u+1/u),\quad k=(d/2)(1/v+1/c),$$
$$d/u=r-w,\quad w=(d/2)(u-1/u),\quad k'=(d/2)(1/v-1/c).$$

The resulting integral

$$I_1 = e^{-\gamma(t_1+x_1/v)}/v \int_{u_2}^{u_1} e^{k\gamma u - k'\gamma/u}\, du/u$$

can be expanded in powers of k' if this is small compared to k, from the formulae

$$\int_{u_2}^{u_1} e^{k\gamma u - k'\gamma/u}\, du/u = \sum_{n=0}^{\infty} (-k'\gamma)^n/n! \int_{u_2}^{u_1} e^{k\gamma u}\, du/u^{n+1},$$

$$\int_{u_2}^{u_1} e^{\alpha u}\, u^{-n-1}\, du = F_n(u_2) - F_n(u_1) + \alpha^n/n! \int_{u_2}^{u_1} e^{\alpha u}\, du/u,$$

$$F_n(u) = e^{\alpha u}\,[u^{-n}/n + \alpha u^{-n+1}/n(n-1) + \cdots + \alpha^{n-1}u^{-1}/n!]$$

$$\int_{u_2}^{u_1} e^{k\gamma u}\, du/u = Ei(k\gamma u_1) - Ei(k\gamma u_2).$$

The function Eix can be expressed in terms of the logarithmic integral,* and for small u,

$$Ei(u) = 0.5772 + \log u + u + u^2/2!\,2 + u^3/3!\,3 + u^4/4!\,4 + \cdots$$

while if u is large,

$$Ei(u) = e^u/u\,[1 + 1!/u + 2!/u^2 + 3!/u^3 + \cdots].$$

The function V_1 is obtained from I_1 by differentiation, $\bar{t}_1$ not being differentiated, and the integral with v replaced by $-v$:

$$I_2 = -e^{-\gamma(t_1-x_1/v)}/v \int_{1/\bar{u}_1}^{1/\bar{u}_2} e^{k\gamma\bar{u} - k'\gamma/\bar{u}}\, d\bar{u}/\bar{u}.$$

Let $v = c$, whence $k' = 0$, $k = d/c$. The integral limits are,

$$du_1 = x_1 + ct_1, \qquad du_2 = (d^2 + x_1^2)^{1/2} + x_1,$$
$$d/\bar{u}_1 = ct_1 - x_1, \qquad d/\bar{u}_2 = (d^2 + x_1^2)^{1/2} - x_1.$$

The terms containing the integrals in the expression for V_1 reduce to $Q_0/4\pi$ multiplied by

$$\begin{aligned} v(\partial/\partial x_1)(I_1 - I_2) = &-2e^{-\gamma(t_1-r_2/c)}/r_2 + 1/r(w_1) + 1/r(\bar{w}_1) \\ &-(\gamma/c)e^{-\gamma(t_1+x_1/c)}[Ei(k\gamma u_1) - Ei(k\gamma u_2)] \\ &+(\gamma/c)e^{-\gamma(t_1-x_1/c)}[Ei(k\gamma/\bar{u}_2) - Ei(k\gamma/\bar{u}_1)]. \end{aligned}$$

* Jahnke-Emde, *Funktionentafeln*, p. 19.

Now $Ei(k\gamma u_1)$, $Ei(k\gamma/\bar{u}_1)$ are independent of d, and cancel in the difference $V_1 - V_2$. Consequently if $R = G = 0$, $(LC)^{-1/2} = c$, the expression for the voltage V becomes,

$$
\begin{aligned}
(3.1)\quad V(x_1, t_1) = &-2\Phi(x_1, t_1) + (Q_0\gamma/4\pi c)e^{-\gamma(t_1+x_1/c)} \\
&\times \{Ei[(\gamma/c)(x_1 + [d^2 + x_1^2]^{1/2})] - Ei[(\gamma/c)(x_1 + [d'^2 + x_1^2]^{1/2})]\} \\
&+ (Q_0\gamma/4\pi c)e^{-\gamma(t_1-x_1/c)} \\
&\times \{Ei[(\gamma/c)([d^2 + x_1^2]^{1/2} - x_1)] - Ei[(\gamma/c)([d'^2 + x_1^2]^{1/2} - x_1)]\}.
\end{aligned}
$$

For $x_1/d \gg 1$,

$$
\begin{aligned}
&Ei[(\gamma/c)([d^2 + x_1^2]^{1/2} + x_1)] \\
&\qquad \sim Ei[(\gamma/c)([d'^2 + x_1^2]^{1/2} + x_1)] \sim e^{2\gamma x_1/c}/2\gamma x_1,
\end{aligned}
$$

hence the first bracket contributes to V a term of the form

$$(Q_0/8\pi x_1^2)e^{-\gamma(t_1-x_1/c)}(\cdot\ \cdot\ \cdot).$$

The second bracket contributes asymptotically

$$(Q_0\gamma/4\pi c)e^{-\gamma(t_1-x_1/c)}\{\log[(d^2 + x_1^2)^{1/2} - x_1] - \log[(d'^2 + x_1^2)^{1/2} - x_1]\}$$

and from the definition $d^2 = (\bar{z} - h)^2 + \bar{y}^2$, $d'^2 = (\bar{z} + h)^2 + \bar{y}^2$, with the result

$$(3.2)\quad V(x_1, t_1) \sim V' = -(Q_0\gamma/\pi c)e^{-\gamma(t_1-x_1/c)}\, h\bar{z}/(\bar{z}^2 + h^2 + \bar{y}^2),\ t_1 > x_1/c.$$

The expression V' represents a voltage wave which jumps from 0 to a maximum at the time $t_1 = x_1/c$, and falls to e^{-1} times this maximum in the interval $\Delta t = 1/\gamma$.

If the charge Q is distributed uniformly over a rectangle $a \leq x_0 \leq b$, $c \leq \bar{y} \leq d$ at the time $t = 0$, let

$$Q_0 = \sigma(b - a)(d - c), \quad \alpha = -\gamma\sigma h\bar{z}/\pi c(h^2 + \bar{y}^2 + \bar{z}^2).$$

A charge at $(x_0, \bar{y}, \bar{z})$ contributes to the voltage wave at (x_1, t_1) only if $a \leq x_0 \leq b$, $x_1 - x_0 \geq ct_1$, replacing x_1 by $x_1 - x_0$ in V' and integrating, the three cases result:

$$
\begin{aligned}
&(1)\quad ct_1 < x_1 - b, \qquad V' = 0. \\
(3.3)\quad &(2)\quad x_1 - a > ct_1 > x_1 - b, \\
&\qquad V' = \int_{x_1-ct_1}^{b} \alpha e^{-\gamma(t_1-x_1/c)-\gamma x_0/c}\, dx_0 = (\alpha c/\gamma)[1 - e^{-\gamma[t_1-(x_1-b)/c]}]. \\
&(3)\quad ct_1 > x_1 - a, \qquad V' = (\alpha c/\gamma)[e^{-\gamma a/c} - e^{-\gamma b/c}]\, e^{-\gamma(t_1-x_1/c)}.
\end{aligned}
$$

Integrating with respect to $\bar{y}$ and neglecting h compared to $\bar{z}$,

$$\int_c^d (\alpha c/\gamma)\,d\bar{y} \doteq -(\sigma h/\pi)\tan^{-1}[(d-c)/(\bar{z}+cd/\bar{z})] = A.$$

The maximum voltage results for $t_1 = (x_1 - a)/c$:

$$V'_{max} = A(1 - e^{-\gamma(b-a)/c}).$$

and the time required to reach this maximum is $(b-a)/c$.

If the voltage due to a point charge is assumed to be that due to a redistribution of the charge on the line at $t = 0$, we have

$$\begin{aligned} V &= -\Phi_0(x_1 - vt_1) - \Phi_0(x_1 + vt_1). \\ V_{max} &\doteq -\Phi_0(0) = -(Q_0/4\pi)(1/d - 1/d'). \\ &\doteq -(Q_0/4\pi)2\bar{z}h/(\bar{z}^2 + h^2 + \bar{y}^2)^{3/2}. \end{aligned}$$

From (3.2),

$$V'_{max}/V_{max} \doteq (2\gamma/c)(\bar{z}^2 + h^2 + \bar{y}^2)^{1/2} \doteq 2\gamma\tau,$$

if τ is the time required for a discontinuity to travel from the cloud to the line.

HELICES IN EUCLIDEAN *N*-SPACE.

By J. H. BUTCHART.

Introduction. A necessary and sufficient condition that a curve $x_i(s)$ in euclidean three-space E_3 be a helix is that the rank of the determinant $| x_i^{(j)} |$ $(i = 1, 2, 3;\ j = 2, 3, 4)$ be two. This paper generalizes this for E_n and points out some properties of helices and related curves termed pseudo-helices.

PART I. FORMULAS FOR A GENERAL CURVE.

Let the curve C be given by $x_i = x_i(s)$ $(i = 1, \cdots, n)$, where s is the arc-length. Then $x_i^{(1)}$, where the differentiation is with respect to the arc, is the unit tangent vector. We assume that C is not contained in any hyper-plane, so that the n vectors $x_i^{(j)}$ $(j = 1, \cdots, n)$ are independent. Following the notation of Eisenhart,* we set up the quantities

$$\sum_{i=1}^{n} x_i^{(p)} x_i^{(q)} = b_p{}^q = b_q{}^p \tag{1}$$

and we denote by b_p the determinant

$$b_p = | b_\alpha{}^\beta | \qquad (\alpha, \beta = 1, \cdots, p). \tag{2}$$

These equations do not define b_0, which we shall take to be unity. The equations

$$\lambda_p|^i = (b_p/b_{p-1})^{1/2} \sum_{\alpha=1}^{p} x_i^{(\alpha)} B_p{}^{(\alpha)} \qquad (p = 1, \cdots, n), \tag{3}$$

where $B_p{}^\alpha$ is the cofactor of $b_p{}^\alpha$ in b_p divided by b_p, define an orthogonal ennuple of unit vectors, which may be called the tangent and principal normals of C. Blaschke gives the curvatures in the formula †

$$1/\rho_p = (b_{p-1} b_{p+1})^{1/2} / b_p \qquad (p = 1, \cdots, n-1). \tag{4}$$

He arrives at this formula by generalizing the Frenet-Serret equations to

$$d\lambda^i{}_p|/ds = -\lambda^i{}_{p-1}|/\rho_{p-1} + \lambda^i{}_{p+1}|/\rho \qquad (p = 1, \cdots, n) \tag{5}$$

with the understanding that $1/\rho_0 = 1/\rho_n = 0$.

* L. P. Eisenhart, *Riemannian Geometry*, Princeton University Press, 1926, p. 104.

† Wilhelm Blaschke, *Mathematische Zeitschrift*, Vol. 6 (1920), pp. 94-99.

It is clear that $b_1 = b_1{}^1 = 1$, and from (4) that $b_2 = 1/\rho_1{}^2$. By induction we obtain the formulas *

$$(6) \qquad b_p = 1/\rho^2{}_{p-1}\rho^4{}_{p-2}\cdots\rho_1{}^{2p-2} \qquad (p=2,\cdots,n).$$

The vectors $\lambda^i{}_{p|}$ were defined as linear combinations of $x_i{}^{(j)}$. We now express $x_i{}^{(j)}$ as

$$(7) \qquad x_i{}^{(j)} = \sum_{\alpha=1}^{n} D_\alpha{}^j \lambda^i{}_{\alpha|}$$

and by the Frenet equations (5) find the reduction formulas

$$(8) \qquad D_\alpha{}^{j+1} = D^j{}_{\alpha-1}/\rho_{\alpha-1} + D_\alpha{}^j - D^j{}_{\alpha+1}/\rho_\alpha$$
$$(j=1,\cdots,n-1;\ 1/\rho_0 = 1/\rho_n = 0).$$

We note that $D_\alpha{}^1 = \delta_\alpha{}^1$ and that $D_\alpha{}^j = 0$ $(\alpha > j)$ and that

$$D_j{}^j = 1/\rho_{j-1}\rho_{j-2}\cdots\rho_1 \qquad (j=2,\cdots,n).$$

We now introduce the quantity B_p, which we define to be the cofactor of $b_1{}^1$ in b_p divided by b_p. Since it is an invariant, we may evaluate it in the cartesian system where the tangent and normals at an arbitrary point of C are taken as the axes. At the origin, $\lambda^i{}_{\alpha|} = \delta_\alpha{}^i$, and $x_i{}^{(j)} = D_i{}^{(j)}$. Since $b_n = |\, x_i{}^{(j)}\,|^2$ $(i,j=1,\cdots,n)$. We can express b_pB_p as the sum of squares of all the $(p-1)$-rowed determinaints which can be formed from the second to the p-th rows of $|\,D_i{}^j\,|$. Expand these determinants by the minors of the elements in the last column, use $b_p = (D_1{}^1D_2{}^2\cdots D_p{}^p)^2$, and the new symbol $D_p = |\,D_\alpha{}^\beta\,|$ $(\alpha = 1,\cdots,p-1;\beta=2,\cdots,p)$. We get

$$B_p = [D_p/D_1{}^1\cdots D_p{}^p]^2 + B_{p-1} = T_p{}^2 + B_{p-1},$$

where T_p is an abbreviation. Continuing this we reach

$$(9) \qquad B_p = \sum_{i=1}^{p} [D_i/D_1{}^1\cdots D_i{}^i]^2 = \sum_{i=1}^{p} T_i{}^2$$
$$(p=1,\cdots,n),$$

in which D_1 is defined as unity. From an induction proof, an evaluation of D_p is

$$(10) \qquad D_p = \sum_{i=1}^{p-1} (-1)^{p-1-i} D_i{}^p D^{p-1}_{p-1} D^{p-2}_{p-2} \cdots D^{i+1}_{i+1} D_i \qquad (p=2,\cdots,n).$$

To illustrate,

$$(11) \quad B_5 = 1 + \rho_2{}^2(1/\rho_1)^2 + \rho_3{}^2(\rho_2/\rho_1)^{(1)^2} + \rho_4{}^2[\rho_2/\rho_3\rho_1 + \{\rho_3(\rho_2/\rho_1)^{(1)}\}^{(1)}]^2.$$

---

* Duschek-Mayer, *Differentialgeometrie*, Vol. 2 (1930), p. 76, B. G. Teubner, Leipzig.

A neater construction for B_p may be given which holds except for special cases. Differentiate D_p by columns and expand the resulting determinants by minors of the differentiated columns. Use (10) and get

$$D_p^{(1)} = (D_p^p/\rho_{p-1})D_{p-1} + (D_p^{p+1}/D_p^p)D_p - D_{p+1}/D_p^p \tag{12}$$
$$(p = 1, \cdots, n-1;\ 1/\rho_0 = 0).$$

Using this and the value of $D_a^{j(1)}$ from (8), we derive

$$T_p = -\rho_{p-1}[2T_{p-1}]^{-1} \frac{d}{ds} B_{p-1}. \tag{13}$$

The foregoing evaluation of B_p proceeded from the assumption that C does not lie in any hyperplane. If instead we take C in an E_{n-1}, the definition of B_n fails since $b_n = 0$. Since $D_n^n = 0$, $b_n B_n = (D_n)^2$, and this is not in general equal to zero. The quantity B_n/ρ^2_{n-1} is likewise well defined.

Part II. Definition and Characterization of a Helix.

Let C represent the helix $x_i = \phi_i(\sigma)$, $x_n = \sigma \cot \theta$, where σ is the arc of the directrix $\bar{C}$, which differs from C only in $x_n = 0$. The relation $\sigma = s \sin \theta$ between the arcs of $\bar{C}$ and C is easily obtained by differentiating the coördinates of C with respect to s and summing their squares. Hence C is $x_i = \psi_i(s)$, $x_n = s \cos \theta$. To derive a necessary and sufficient condition that a curve be a helix, we shall express the $n-1$ curvatures of the helix in terms of the $n-2$ curvatures of the directrix. The quantities b_p^q, $\bar{b}_p^q$ for the helix and directrix, respectively, are joined by the relations

$$\begin{aligned} b_1^1 &= (\bar{b}_1^1 + \cot^2\theta)\sin^2\theta \\ b_p^q &= \bar{b}_p^q \sin^{p+q}\theta \qquad (p, q \text{ not both } 1). \end{aligned} \tag{14}$$

Using these in b_p, we obtain

$$b_p = \bar{b}_p(1 + \bar{B}_p \cot^2\theta)\sin^{p(p+1)}\theta \qquad (p = 1, \cdots, n).$$

Combine these with (4) and get

$$\frac{1}{\rho_p} = \frac{\sin\theta}{\bar{\rho}_p} \frac{\sqrt{(1 + \bar{B}_{p-1}\cot^2\theta)(1 + \bar{B}_{p+1}\cot^2\theta)}}{1 + \bar{B}_p \cot^2\theta} \tag{15}$$
$$(p = 1, \cdots, n-1).$$

The relations (14) may be solved for $\bar{b}_p^q$ in terms of b_p^q, and these are used to derive

$$\bar{b}_p = b_p(1 - B_p \cos^2\theta)\csc^{p(p+1)}\theta \qquad (p = 1, \cdots, n), \tag{16}$$

and

$$\frac{1}{\bar{\rho}_p} = \frac{\csc\theta}{\rho_p} \frac{\sqrt{(1 - B_{p-1}\cos^2\theta)(1 - B_{p+1}\cos^2\theta)}}{1 - B_p\cos^2\theta} \qquad (p = 1, \cdots, n-1). \tag{17}$$

Since $\bar{C}$ is a hyperplane curve, from (16) $B_n = \sec^2\theta$ is a necessary condition for C to be a helix. For the sufficiency proof, we need to use the further condition that $T_n \neq 0$. We construct a curve by (17) and notice that its helix with curvatures given by (15) is congruent to the original curve. By (13) the condition $B_n = \text{const.}$, $T_n \neq 0$ is equivalent to $D_{n+1} = 0$, $D_n \neq 0$, so we have the theorem:

A necessary and sufficient condition that a curve in euclidean n-space be a helix is that the rank of the determinant $| x_i^{(j)} |$ $(i = 1, \cdots, n;\ j = 2, \cdots, n+1)$ *be* $n - 1$.

Part III. Geometric Properties.

Pseudo-helices. If for a curve C, $| x_i^{(j)} | \begin{pmatrix} i = 1, \cdots, n; \\ j = 2, \cdots, n+1 \end{pmatrix}$ is of rank $n - k$, we may call it a pseudo-helix of class k. From (17) and (15) we have the theorem:

A necessary and sufficient condition that a curve be a pseudo-helix of class k is that there exist a helix H in E_{n-k} in arc correspondence with C such that at corresponding points of H and C the curvatures of H are identical with the first $n - k - 1$ curvatures of C.

From (11), a curve in E_4 is a pseudo-helix if and only if ρ_2/ρ_1 be constant. If we call the system of planes, which are the ultimate intersections of neighboring three-spaces taken orthogonal to the principal normal, the rectifying three-space developable, we may easily show that:

A pseudo-helix in E_4 cuts the generating planes of its rectifying three-space developable under the constant angle $\phi = \tan^{-1}(\rho_2/\rho_1)$.

If the first q pairs of consecutive curvatures have constant ratios, i. e., $\rho_{2p}/\rho_{2p-1} = k_p$ $(p = 1, \cdots, q)$, then from (10) and (8) by an induction proof we can show that the quantities D_p are alternately constant and zero. Hence:

If $\rho_{2p}/\rho_{2p-1} = k_p$ $(p = 1, \cdots, q)$, where k_p is constant, then the curve is a helix if it lies in E_{2q+1} and is a pseudo-helix if it lies in E_n.

As a corollary we may state:

If a curve all of whose curvatures are constant lies in a space of an even

number of dimensions, it is a pseudo-helix; and if it lies in an odd space, it is a helix whose directrix is a pseudo-helix.

This may be proved independently by (8) and the definition of D_p.

Other geometric properties whose proofs are fairly direct are:

A helix whose directrix is a helix is contained in a space of the same number of dimensions as the space in which the directrix lies.

If a helix is of the same number of dimensions as its directrix, this directrix is either a helix or a pseudo-helix.

Each of the ∞^2 helices H on a directrix D in E_{n-2} has a directrix C in every cylinder on D normal to E_{n-2}, and these curves C are themselves helices of D.

The third curvature of a helix on a curved geodesic of a developable surface in E_3 is unchanged by any deformation of the developable which carries generators into generators.

CHARACTERIZATIONS OF CERTAIN CURVES BY CONTINUOUS FUNCTIONS DEFINED UPON THEM.

By Gordon T. Whyburn.

Cech has shown (*Fundamenta Mathematicae,* Vol. 18) that any compact continuum M upon which there can be defined a real, continuous function which is not constant on any infinite subset of M is a particular kind of regular curve.* Mazurkiewicz (*Fundamenta Mathematicae,* Vol. 18) has given necessary and sufficient conditions in order for a given acyclic locally connected continuum (dendrite) to have the property of admitting such a function to be defined upon it. In the present paper it will be shown that by lightening the restrictions on the function to varying degrees, regular, rational and 1-dimensional curves in the Menger-Urysohn sense may be characterized among the compact metric continua by the admission of such functions to be defined upon them.* Our results are embodied in the following proposition.

Theorem. *In order that the compact continuum M be a* $\left\{\begin{matrix}\textit{regular}\\ \textit{rational}\\ \textit{1-dimensional}\end{matrix}\right\}$ *curve it is necessary and sufficient that there exist a real, continuous function $f(p)$ defined on M which is not constant on any subcontinuum of M and which takes each of an everywhere dense set of its values only on the points of a* $\left\{\begin{matrix}\textit{finite}\\ \textit{countable}\\ \textit{0-dimensional}\end{matrix}\right\}$ *set.*

We shall first show that the condition is sufficient. For simplicity of expression we shall use the terms R_1-curve, R_2-curve, R_3-curve to mean regular curve, rational curve, and 1-dimensional curve, respectively, and the terms T_1-set, T_2-set, T_3-set to mean finite set, countable set, and 0-dimensional set, respectively.

Now let M be any compact continuum upon which there exists a real, continuous function $f(p)$ which is not constant on any subcontinuum of M and such that an everywhere dense set E of its values exists such that for each number e of E the set M_e where $f(p) = e$ is a T_i-set $(i = 1, 2, 3)$. We shall prove that M is an R_i-curve. Suppose this is not true. Then † there

---

* We employ the usual terminology and symbolism of Point Set Theory. For definitions of the terms used see, for example, Menger, *Kurventheorie,* B. G. Teubner, 1932.

† See Menger, *Kurventheorie,* pp. 128-133, where references in this connection also will be found to Hurewicz, Urysohn, and others.

exists a point x of M and a non-degenerate subcontinuum N of M containing x and such that no point of $N - x$ can be separated from x in M by any T_i-set. By hypothesis there exists a point y of $N - x$ such that $f(x) \neq f(y)$. But if e is a number of the set E which is between $f(x)$ and $f(y)$, then the set M_e of all points p such that $f(p) = e$ is a T_i-set and since f is continuous, it follows at once that M_e separates x and y in M; for if M_1 and M_2 are the sets of all points p of M such that $f(p) < e$ and $f(p) > e$, respectively, clearly M_1 and M_2 are mutually separated, one contains x and the other y, and $M_1 + M_2 = M - M_e$. Thus the supposition that M is not an R_i-curve leads to a contradiction.

Our proof for the necessity of the condition will be based on the following

LEMMA. *If A and B are disjoint, closed T_i-subsets $(i = 1, 2, 3)$ of a compact R_i-set* K and ϵ is any positive number, then there exist closed, disjoint, T_i-subsets $A = X_0, X_1, X_2, X_3, X_4 = B$ and closel subsets K_0, K_1, K_2, K_3 of K such that*

(i) $\sum_0^3 K_n = K$; $K_{n-1} \cdot K_n = X_n$; $K_n \cdot K_m = 0$ if $|m - n| > 1$;

(ii) *if x and y are any two points of K_0 and K_3, respectively then* $\rho(A, x) < \epsilon > \rho(B, y)$;

(iii) *every subcontinuum of K of diameter $> \epsilon$ intersects $\sum_0^4 X_n$.*

By virtue of the general decomposition theorem for curves † we can write $K = H_1 + H_2 + \cdots + H_k$, where each H_m is closed, $\delta(H_n) < \min [\epsilon, 1/8 \rho(A, B)]$, $H_m \cdot H_n$ is a T_i-set, $H_m \cdot H_n \cdot (A + B) = 0$, and $H_m \cdot H_n \cdot H_r = 0$ for $m \neq n \neq r \neq m$. Let U, W, V be the sum of all those sets H_n which contain at least one point of A, which contain at least one point of B, and which contain no point of $A + B$, respectively. We can now define our required closed sets X_n and K_n as follows:

$X_0 = A$,
$X_1 = U \cdot V + \sum H_m \cdot H_n$ where $H_m + H_n \subset U$, $m \neq n$,
$X_2 = \sum H_m \cdot H_n$, where $H_m + H_n \subset V$,
$X_3 = V \cdot W + \sum H_m \cdot H_n$, where $H_m + H_n \subset W$,
$X_4 = B$;
$K_0 = U$,
$K_1 = X_1 + X_2 + \sum H_n$ such that $H_n \subset V$ and $H_n \cdot U \neq 0$
$K_2 = X_2 + X_3 + \sum H_n$ such that $H_n \subset V$ and $H_n \cdot U = 0$
$K_3 = W$.

* That is, a set each point of which is contained in arbitrarily small neighborhoods whose boundaries are T_i-sets.

† See, for example, Menger, *Kurventheorie*, p. 183.

It can be verified at once that these sets satisfy all the conditions required by the lemma.

Now, to prove that the condition in our theorem is necessary, let M be any R_i-curve and let a and b be any two distinct points of M. Let us take $\epsilon = 1$ and, using $M = K$, $a = A$, $b = B$, obtain the sets $X_0, \cdots, X_4$, $K_0, \cdots, K_3$ as in the lemma. Set $X_m = X(m/4)$, $(0 \leq m \leq 4)$ and $K_m = K(m/4)$, $(0 \leq m < 4)$. Now supposing that for any integer $n \geq 1$, the sets $X(m/4^n)$, $(0 \leq m \leq 4^n)$ and $K(m/4^n)$ $(0 \leq m < 4^n)$ have already been defined (and indeed they have been defined for $n = 1$), we define the set $X(m/4^{n+1})$ and $K(m/4^{n+1})$ as follows: Take $\epsilon = 1/(n+1)$ and, for any m, $0 \leq m < 4^n$, using $K(m/4^n) = K$, $X(m/4^n) = A$, $X[(m+1)/4^n] = B$, we obtain the sets $X_0, \cdots, X_4$, $K_0, \cdots, K_3$ as in the lemma; then set $X_k = X[(4m+k)/4^{n+1}]$, $(0 \leq k \leq 4)$, and $K_k = K[(4m+k)/4^{n+1}]$, $(0 \leq k < 4)$.

We thus obtain, for each positive integer n, a collection of $4^n + 1$ disjoint closed sets $X(m/4^n)$, $(0 \leq m \leq 4^n)$, and a collection of 4^n closed sets $K(m/4^n)$, $(0 \leq m < 4^n)$, having the following properties:

(1) $\sum_{m=0}^{4^n-1} K(m/4^n) = M$; $K[(m-1)/4^n] \cdot K(m/4^n) = X(m/4^n)$; $K(m/4^n) \cdot K(j/4^n) = 0$ if $|m-j| > 1$; $X(0/4^n) = a$, $X(4^n/4^n) = b$;

(2) if an index $m/4^n$ is expressible in the form $j/4^k$, $k < n$, $0 \leq j \leq 4^k$, then: $X(m/4^n) = X(j/4^k)$, $K(m/4^n) \subset K(j/4^k)$ and every point of $K[(m-1)/4^n] + K(m/4^n)$ is at a distance $< 1/n$ from $X(m/4^n)$.

(3) if $k > n$ and $K(m/4^n) \supset K(j/4^k)$, then $m/4^n \leq j/4^k \leq (m+1)/4^n$.

(4) every subcontinuum of M of diameter $> 1/n$ intersects at least one of the set $X(m/4^n)$.

We now define a function $f(p)$ on M as follows: If for some n, p belongs to $X(m/4^n)$, then let $f(p) = m/4^n$. If p belongs to no set $X(m/4^n)$, then for each n there exists exactly one integer j_n such that p belongs to the set $K(j_n/4^n)$. In this case since, for each n, $K(j_n/4^n) \subset K(j_{n-1}/4^{n-1})$ it follows by (3) that the sequence of numbers $[j_n/4^n]$ converges, and we set $f(p) = \lim_{n\to\infty} j_n/4^n$.

Then $f(p)$ satisfies all the conditions required for our theorem. Obviously $f(p)$ is real and $0 \leq f(p) \leq 1$. Furthermore, $f(p)$ is continuous. For let x be any point of M, let ϵ be any positive number and let us choose n so large that $1/4^{n-1} < \epsilon$. Let m be the largest integer such that x belongs to $K(m/4^n)$. Then either x belongs to no other set $K(j/4^n)$ or it belongs also to $K[(m-1)/4^n]$ but to no other. In either case let V and W respectively

denote the sum of all those sets $K(j/4^n)$ which do and which do not contain x. Then since W is closed, there exists a neighborhood U of x such that $U \cdot W = 0$. Then for any point p of $U \cdot M$, we have $p \subset V$ and hence, by (3), $(m-1)/4^n \leq f(p) \leq (m+1)/4^n$; and since both p and x belong to V, this gives $|f(p) - f(x)| < \epsilon$, which proves f continuous.

Now let E be the set of all numbers on $(0, 1)$ of the form $m/4^n$, $(0 \leq m \leq 4^n)$, and let $e = m/4^n$ be any number of E. Then if x is any point of the set $X(m/4^n)$, we have $f(x) = e$; and if x does not belong to $X(m/4^n)$ and is at a distance d from it, then by (2) we can choose $k > n$ so large that $1/k < d$ and x will not belong to the set $K[(j-1)/4^k] + K(j/4^k)$, where $j/4^k = m/4^n$. Whence by (3), we have either $f(x) \geq (j+1)/4^k$ or $f(x) \leq (j-1)/4^k$, either of which gives $f(x) \neq e$. Thus $f(p) = e$ if and only if $p \subset X(e)$; and since $X(e)$ is a T_i-set, we have an everywhere dense set E of values of f each of which it takes only on a T_i-set. Finally, since no non-degenerate subcontinuum of M is a T_i-set, for any i, and since by (4) every such continuum in M must intersect some set $X(e)$, it follows that f cannot be constant on any subcontinuum of M. This completes the proof.

It will be noted that in the third * of the three parts to our general theorem, the final condition on the function is entirely superfluous, because the condition that f be not constant on any subcontinuum of M implies that for *every* value e of f, the set M_e of points p such that $f(p) = e$ is 0-dimensional. Thus we have the following

COROLLARY. *In order that the compact continuum M be 1-dimensional (i. e., be a "curve" in the Menger-Urysohn sense) it is necessary and sufficient that there exist a real continuous function defined on M which is not constant on any subcontinuum of M.*

THE JOHNS HOPKINS UNIVERSITY.

* A similar remark would not apply to either of the other two parts, however. For if f takes the value e, $0 < e < 1$, on a set M_e of power less than that of the continuum, then M_e contains a local separating point of M. Thus if M is any continuum having at most a countable number of local separating points, e. g., the Sierpinski triangle curve, then any real continuous function on M must take all but a countable number of its values on a set having the power of the continuum.

ON UNICOHERENCY ABOUT A SIMPLE CLOSED CURVE.*

By W. A. WILSON.

1. The subject matter of this article is a discussion of a property of point-sets which is closely related to the properties of being connected and of being a unicoherent continuum, and is in a sense a generalization of both these properties. A unicoherent continuum is defined by Kuratowski † as one which cannot be expressed as the union of two continua whose divisor is not connected, and this definition is in common usage. The analogy of the definition to that of a continuum is apparent. However, a set that is not connected may be connected between some pair of its points. We therefore propose a corresponding definition of unicoherency.

A set M in a metric space is unicoherent about the simple closed curve J if, for every decomposition of J into closed arcs h and k by points a and b and for every decomposition of M into relatively closed sets H and K such that $h \subset H$, $k \subset K$, and $h \cdot K = k \cdot H = a + b$, there is a component of $H \cdot K$ containing both a and b. (If M is closed or is the space itself, the word "relatively" is to be omitted.)

It is this extension of the idea of unicoherent continuum that is to be discussed. That this may be regarded also as a natural extension of the definition of connectivity between two points is apparent when we recall that in a one-dimensional Euclidean space an interval is a "sphere" and two points form its frontier or the "surface of the sphere." ‡ The property of a set being unicoherent about a simple closed curve is also useful in formulating an intrinsic definition of a two-dimensional simplex, as will be seen later. Readers of H. Whitney's recent article on this subject § will note the marked resemblance to the property of a simple closed curve being homologous to zero in a closed set, which is the basis of his article. In fact, the present article grew out of a search for purely point-set properties equivalent to Whitney's definition.

* Presented to the Society, October, 1932.

† C. Kuratowski, "Sur la structure des frontières communes à deux régions," *Fundamenta Mathematicae*, Vol. 12, p. 24.

‡ The further generalizations suggested by this analogy will of course occur to the reader.

§ H. Whitney, "A characterization of the simple 2-cell," *Transactions of the American Mathematical Society*, Vol. 35.

The following four sections deal with general properties of metric or compact metric spaces unicoherent about a simple closed curve. Sections 6-8 contain theorems useful in the study of sets irreducible with respect to these properties. The remainder of the article gives some of the consequences of imposing the condition of local connectivity and includes the intrinsic definition of the two-dimensional simplex already mentioned.

2. The definition given in the previous section is equivalent to the following somewhat more compact one: *A metric space Z is unicoherent about the simple closed curve J if, for every partition of J into open arcs λ and μ by points a and b and for every decomposition of Z into closed sets H and K such that $\lambda \cdot K = \mu \cdot H = 0$, there is a component of $H \cdot K$ containing both a and b.* A similar definition holds for a set M imbedded in a metric space.

It is well known that, if A and B are separated sets (i. e., $\bar{A} \cdot B + A \cdot \bar{B} = 0$) in a metric space Z, then there are separated regions R and S such that $A \subset R$ and $B \subset S$. Hence Z is the union of closed sets H and K such that $A \subset H$, $B \subset K$, $A \cdot K = 0$, and $B \cdot H = 0$. Since the open arcs λ and μ in the above definition are separated sets, the definition is never trivially satisfied by there being no decomposition of Z into closed sets H and K such that $\lambda \cdot K = \mu \cdot H = 0$.

A metric space which is unicoherent about a simple closed curve need not be connected nor, if it is connected, is it necessarily a unicoherent continuum. The first statement is obvious; two examples of the second will be given. The first is the plane set M consisting of two externally tangent circumferences J and K and the interior of one of them, say of J. That M is unicoherent about J is a consequence of the Phragmen-Brouwer theorem; it is clearly not a unicoherent continuum. The second example is a hemispherical surface cut off by a circumference J which has two points not on J pinched together to form what may be called a double point. This is unicoherent about J (See § 5), but it is not a unicoherent continuum. For it is the union of two continua whose divisor is the double point and a properly drawn arc joining two points of J.

On the other hand, a unicoherent continuum may fail to be unicoherent about some simple closed curve contained in it. For example, let Z be the sum of a circumference J and a spiral approaching J as a limit. We note that here Z is not locally connected.

3. The opposite of the definitions of unicoherency merely says that, if the metric space Z is not unicoherent about the simple closed curve J, there is *some* pair of points a and b dividing J into open arcs λ and μ and *some*

decomposition of Z into closed sets H and K such that $\lambda \cdot K = \mu \cdot H = 0$ and no component of $H \cdot K$ joins a and b. We have, however, a stronger statement for compact metric spaces.

THEOREM. *Let the compact metric space Z contain the simple closed curve J and be not unicoherent about J. Let a and b be any two points which divide J into open arcs λ and μ. Then Z is the union of closed sets H and K such that $\lambda \cdot K = \mu \cdot H = 0$ and no component of $H \cdot K$ contains a and b.*

Proof. By the definition of unicoherency there is some pair of points c and d dividing J into open arcs α and β, and some decomposition of Z into closed sets P and Q such that $\beta \cdot P = \alpha \cdot Q = 0$ and no component of $P \cdot Q$ contains c and d. If $a + b = c + d$, this is the required decomposition.

In the contrary event suppose that a lies on the arc α. Then a divides α into open arcs γ, whose end-points are a and c, and δ, whose end-points are a and d. By the previous paragraph $P \cdot Q = C + D$, where C and D are disjoint closed sets containing c and d, respectively. Then $C + \gamma$ and $D + \delta$ are separated sets, and P is the union of closed sets R and S, such that $R \cdot (D + \delta) = S \cdot (C + \gamma) = 0$. Now a and c divide J into the open arcs γ and $\beta + d + \delta$, and R and $Q + S$ are closed sets such that $R \cdot (\beta + d + \delta) = (S + Q) \cdot \gamma = 0$. The divisor of R and $Q + S$ is $Q \cdot R + R \cdot S$. Since $Q \cdot R \subseteq C$ and $R \cdot S \cdot C = 0$, no component of $R \cdot (Q + S)$ contains a and c. Thus the conclusion of the theorem is true for the points a and c.

If $b = c$, the theorem is proved. In the contrary event, b lies on one of the open arcs whose end-points are a and c, and we have only to repeat the reasoning of the previous paragraph.

COROLLARY 1. *Let the compact metric space Z contain the simple closed curve J. For Z to be not unicoherent about J it is necessary and sufficient that there be an upper semi-continuous decomposition of Z into disjoint closed sets, each of which contains exactly one point of J.*

COROLLARY 2. *Let the compact metric space Z contain the simple closed curve J. For Z to be not unicoherent about J it is necessary and sufficient that J be the continuous image of Z by a transformation such that, if the point x lies on J, x is the image of itself.*

The first of these corollaries is readily deduced from the theorem and the definition of unicoherency. The second is equivalent to the first by well known theorems on upper semi-continuous decompositions. It follows from

Corollary 2 that the property of Z being not unicoherent about J is equivalent to the property of J being a "retracte" of Z, in the language of K. Borsuk.* A somewhat similar theorem is proved by Borsuk for quasi-peanian spaces (*loc. cit.*).

4. We now proceed to get another equivalent definition of unicoherency. In the statement of the theorem about to follow, it is to be understood that any of the arcs there designated by α, β, λ, and μ may lack one or both end-points, that any one may be a single point, and that either α and β or λ and μ may both be points. The fact that α and β are separated insures that neither λ nor μ is void; and, of course, if α (or β) lacks an end-point, this point lies in λ or μ.

THEOREM. *For the compact metric space Z containing the simple closed curve J to be unicoherent about J it is necessary and sufficient that, if α and β are any separated arcs of J, and λ and μ are the complementary arcs of J, then for every decomposition of Z into closed sets H and K such that $\lambda \cdot K = \mu \cdot H = 0$, some component of $H \cdot K$ contains points of both α and β.*

Proof. That the condition is sufficient follows at once from the theorem in § 3. To show that the condition is necessary is nearly as easy.

Let us suppose that α, β, λ, and μ are as stated in the theorem and that H and K are closed sets such that $Z = H + K$ and $\lambda \cdot K = \mu \cdot H = 0$. If no component of $H \cdot K$ contained points in both α and β, $H \cdot K$ would be the sum of disjoint closed sets A and B such that $\beta \cdot A = \alpha \cdot B = 0$. Let us assume this.

Let c and d be points on λ and μ, respectively. They divide J into open arcs α' and β' containing α and β, respectively. Then $A + \alpha'$ and $B + \beta'$ are separated sets, and so Z is the union of closed sets P and Q such that $P \cdot (B + \beta') = Q(A + \alpha') = 0$. No component of $P \cdot Q$ joins c and d, because it would have to meet $H \cdot K = A + B$. Hence Z is not unicoherent about J. This contradiction shows that our assumption was false, and so the theorem is proved.

5. THEOREM. *In a metric space let M' be a compact set unicoherent about the simple closed curve J' and let M be the continuous image of M' by a transformation such that the correspondence between J' and its image J is a homeomorphism. Then M is unicoherent about J.*

* K. Borsuk, "Quelques théorèmes sur les ensembles unicohérents," *Fundamenta Mathematicae*, Vol. 17, p. 184.

Proof. Let a and b be any points of J, dividing it into the open arcs α and β, and let $M = P + Q$, where P and Q are closed sets, $\alpha \subset P$, $\beta \subset Q$, and $\beta \cdot P = \alpha \cdot Q = 0$. In the correspondence of M' to M let $a' \sim a$, $b' \sim b$, $\alpha' \sim \alpha$, $\beta' \sim \beta$, $P' \sim P$, and $Q' \sim Q$.* Clearly $\alpha' \subset P'$, $\beta' \subset Q'$, and $\beta' \cdot P' = \alpha' \cdot Q' = 0$. On account of the continuity, P' and Q' are closed. Since M' is unicoherent about J', there is a component μ' of $P' \cdot Q'$ joining a' and b'. But then the image μ of μ' is a continuum joining a and b, and $\mu \subset P \cdot Q$. Hence the condition for unicoherency of M about J is satisfied.

6. THEOREM. *Let $\{M_i\}$ be a descending sequence of compact metric sets, each unicoherent about the simple closed curve J, and let M be the divisor of the sequence. Then M is unicoherent about J.*

Proof. If the theorem is not true, there are points a and b on J dividing it into open arcs α and β, and a decomposition of M into closed sets P and Q such that $\alpha \subset P$, $\beta \subset Q$, $\beta \cdot P = \alpha \cdot Q = 0$, and $P \cdot Q$ is the sum of disjoint closed sets A and B containing a and b, respectively.

Let $W_\epsilon(P)$ and $W_\epsilon(Q)$ denote the sets of points of M_1 whose distances from P and Q, respectively, are not more than ϵ. For ϵ small enough $W_\epsilon(P) \times W_\epsilon(Q)$ contains no component joining A and B, in consequence of theorems regarding upper closed limiting sets. For i large enough, $M_i \subset W_\epsilon(M) \subseteq W_\epsilon(P) + W_\epsilon(Q)$. Set $P_i = M_i \cdot W_\epsilon(P)$ and $Q_i = M_i \cdot W_\epsilon(Q)$. Then $\alpha \subset P_i$, $\beta \subset Q_i$, and $P_i \cdot Q_i$ contains no component joining A and B, or, a fortiori, a and b.

Set $P^*_i = M_i - Q_i$ and $Q^*_i = M_i - P_i$. Now $\beta \cdot \bar{P}^*_i = 0$, because $W_\epsilon(\beta) \subset W_\epsilon(Q)$, and so Q_i contains every point of M_i whose distance from β is less than ϵ, whereas $P^*_i \cdot Q_i = 0$. Also $\alpha \cdot \bar{Q}^*_i = 0$. Hence $\alpha + P^*_i$ and $\beta + Q^*_i$ are separated sets and there is a decomposition of M_i into closed sets P'_i and Q'_i such that $\alpha + P^*_i \subset P'_i$, $\beta + Q^*_i \subset Q'_i$, and $P'_i \cdot (\beta + Q^*_i) = Q'_i \cdot (\alpha + P^*_i) = 0$. Now $P'_i \subset P_i$ and $Q'_i \subset Q_i$, whence $P'_i \cdot Q'_i$ has no component joining a and b. This contradicts the hypothesis that M_i is unicoherent about J.

COROLLARY. *A compact metric set M which is unicoherent about a simple closed curve J contains a set N which is irreducible with respect to the property of being closed and unicoherent about J.*

Remarks. It is a simple matter to show that a set irreducible with respect to the property of being closed and unicoherent about a simple closed curve is a continuum. In dealing with such continua we find the same sort

* The correspondences $P' \sim P$ and $Q' \sim Q$ are not in general one to one.

of pitfalls as with continua irreducible between two points and sets irreducibly connected between two points.

We must also guard against confusing the definition of unicoherency here used with that of the unicoherent continuum by Kuratowski. The last example at the end of § 2 is a unicoherent continuum (in the sense of Kuratowski) containing a circumference J, but it contains no set irreducible with respect to the property of being a unicoherent continuum containing J.

It might be thought by analogy with Lennes' definition of the simple arc that a closed set M irreducibly unicoherent about a simple closed curve is a simple 2-cell. That this is false is apparent from the second example given in § 2.

7. Theorem. *Let Z be a compact metric space unicoherent about the simple closed curve J. Let $Z = M + N$, where M and N are closed sets, $J \subset M$, and $M \cdot N$ is a simple arc or a point. Then M is unicoherent about J.*

Proof. Let x be any point of J and assume that M is not unicoherent about J. Then there is an upper semi-continuous decomposition of M into disjoint closed sets $\{M_x\}$ such that each M_x contains exactly one point x, by § 3, Corollary 1. There is also an upper semi-continuous decomposition of N into disjoint closed sets $\{N_y\}$ such that each point y of the arc $M \cdot N$ lies in just one set N_y, as is easily seen by a direct proof or by a theorem of Borsuk.* Let N_x be the union of the sets $\{N_y\}$ corresponding to points $\{y\}$ belonging to M_x. Then $N = \Sigma N_x$ is an upper semi-continuous decomposition of N. Set $Z_x = M_x + N_x$, if $N_x \neq 0$, and $Z_x = M_x$ if $N_x = 0$. Then $Z = \Sigma Z_x$ is an upper semi-continuous decomposition of Z into disjoint closed sets such that each point x of J lies on just one Z_x. Hence Z is not unicoherent about J, by § 3, Corollary 1, which is contrary to the hypothesis. Thus the theorem is proved.

Corollary. *In a compact metric space let M be a set irreducible with respect to the property of being closed and unicoherent about a simple closed curve. Then M has no cut-point.*

For, if c were a cut-point, $M - c$ would be the sum of two separated sets P and Q. Then either $\bar{P}$ or $\bar{Q}$ would contain J and $\bar{P} \cdot \bar{Q} = c$. If $J \subset \bar{P}$, the above theorem shows that $\bar{P}$ would be unicoherent about J, contrary to the hypothesis regarding M.

8. Suppose now that α, β, and γ are disjoint simple open arcs lying in

* K. Borsuk, "Sur les rétractes," *Fundamenta Mathematicae*, Vol. 17, p. 158, § 13.

the compact metric space Z and having common end-points. Suppose also that Z is the union of closed sets M and N, such that $M \cdot N = \bar{\gamma}$, $\alpha \subset M$, and $\beta \subset N$. It would be natural to expect that: (a) if Z is unicoherent about $\bar{\alpha} + \bar{\beta}$, then M and N are unicoherent about $\bar{\alpha} + \bar{\gamma}$ and $\bar{\beta} + \bar{\gamma}$, respectively; and (b) if M and N are unicoherent about $\bar{\alpha} + \bar{\gamma}$ and $\bar{\beta} + \bar{\gamma}$, respectively, then Z is unicoherent about $\bar{\alpha} + \bar{\beta}$. The first of these theorems is readily proved by the same method as that used in the proof of the theorem in § 7 or by means of the theorem of § 3, but the writer has so far been unable to prove the second. The methods used by Borsuk (*loc. cit.*, pp. 190-201) to prove a similar theorem do not seem to be immediately applicable, as we do not have local connectivity.

It will be noted that these two theorems and that in § 7 are the same as Whitney's Lemmas M, N, and O in the article referred to in § 1, provided that for a set to be unicoherent about a simple closed curve is the same as for the simple closed curve to be homologous to zero in the set. A proof of the equivalence of the two properties would of course take care of the second of the above theorems.*

9. We now turn to locally connected compact continua unicoherent about a simple closed curve. The imposition of local connectivity makes a variety of theorems possible, partly on account of the arcwise connectivity of the space, but largely on account of the following property.

THEOREM. *Let Z be a locally connected compact continuum containing connected sets A and B, such that $\bar{A} \cdot B + A \cdot \bar{B} = 0$ and $\bar{A} \cdot \bar{B}$ is totally disconnected. Then Z is the union of locally connected continua M and N such that $B \cdot M = A \cdot N = 0$.*

Proof. Since the sets A and B are separated, Z is the union of closed sets F and G such that $B \cdot F = A \cdot G = 0$. Let R be the component of $Z - G$ containing A; then $\bar{R}$ is a sub-continuum of F containing $\bar{A}$. Let S be the component of $Z - \bar{R}$ containing B and $T = (Z - \bar{R}) - S$. Then $\bar{R} + T$ is a continuum and $\bar{S} \cdot T = 0$. Set $H = \bar{R} + T$ and $K = \bar{S}$. Clearly $Z = H + K$, $\bar{A} \subset H$, $\bar{B} \subset K$, and $A \cdot K = B \cdot H = 0$.

Let $\{\delta_i\}$ be a descending sequence approaching zero. Set $C = \bar{A} \cdot \bar{B}$. Let $V_{\delta_i}(C)$ be the set of points whose distance from C is less than δ_i, and set $T_i = H \cdot V_{\delta_i}(C)$ and $H_1 = H - T_1$. Then H_1 is closed and $\bar{B} \cdot H_1 = 0$.

* Shortly after the submission of this paper to the editors there appeared an article by L. Vietoris in the *Fundamenta Mathematicae* (Vol. 19, pp. 265-273), which contains a theorem including that of § 6 above as a special case, provided that this equivalence is actually true.

Take $\epsilon_1 > 0$ and less than one-third of the distance between B and H_1. Since H_1 is compact, there is a finite set of locally connected continua whose union we call P_1 such that $H_1 \subset P_1 \subset V_{\epsilon_1}(H_1)$.

For each $i > 1$ set $H_i = \overline{T_{i-1} - T_i}$. Then H_i is closed and $\bar{B} \cdot H_i = 0$. Take $\epsilon_i > 0$, less than $\epsilon_{i-1}/2$, and less than one-third of the distance between $\bar{B}$ and H_i. Since H_i is compact, there is a finite set of locally connected continua, each one meeting H_i, whose union we call P_i such that $H_i \subset P_i \subset V_{\epsilon_i}(H_i)$.

Let M_r be the union of the first r sets $\{P_i\}$ and M be the sum of C and $U_1^\infty(M_r)$. It is clear that no M_r meets $\bar{B}$ and that $H \subset M$. Obviously $\epsilon_i \to 0$ and so the points of C are the only improper limiting points of $U_1^\infty(M_r)$. Hence M is closed and $B \cdot M = 0$. Every point of M, except those of C (which lie in H), is on a locally connected sub-continuum of some P_i containing points of H. Hence M is connected. If x is a point of $M - C$, some vicinity of x lies in some M_r; and, since M_r is the union of a finite set of locally connected continua, M is locally connected at x. But then M is necessarily locally connected at the points of C, since the points where a compact continuum is not locally connected cannot form a totally disconnected set.

In like manner we define a locally connected continuum N, such that $\bar{B} \subset K \subset N$ and $A \cdot N = 0$. This completes the proof of the theorem.

10. THEOREM. *Let Z be a locally connected compact continuum containing the simple closed curve J, and let the points a and b divide J into open arcs α and β. Let Z be not unicoherent about J. Then Z is the union of locally connected continua M and N, such that $\alpha \cdot N = \beta \cdot M = 0$ and no component of $M \cdot N$ joins a and b.*

Proof. By § 3, Z is the union of closed sets F and G, such that $\alpha \cdot G = \beta \cdot F = 0$ and no component of $F \cdot G$ joins a and b. If in the first paragraph of the proof in § 9 we replace A and B by α and β, respectively, we see that Z is the union of continua H and K, such that $\bar{\alpha} \subseteq H$, $\bar{\beta} \subseteq K$, and $\alpha \cdot K = \beta \cdot H = 0$. It also follows from this reference that $H \cdot K = \bar{R} \cdot \bar{S} \subset F \cdot G$; consequently no component of $H \cdot K$ joins a and b.

In the rest of the proof in § 9 take $C = a + b$. We then have two locally connected continua M and N whose union is Z, such that $\bar{\alpha} \subset H \subset M \subset V_{\epsilon_1}(H)$, $\bar{\beta} \subset K \subset N \subset V_{\epsilon_1}(K)$, and $\alpha \cdot N = \beta \cdot M = 0$. Since no component of $H \cdot K$ joins a and b, we know from the theory of upper closed limiting sets that no component of $M \cdot N$ joins a and b, provided that ϵ_1 is taken small enough. Hence the theorem is proved.

COROLLARY. *The necessary and sufficient condition for the locally connected compact continuum Z to be unicoherent about the simple closed curve J is that for every pair of points a and b dividing J into open arcs α and β and for every decomposition of Z into locally connected continua M and N such that $\alpha \cdot N = \beta \cdot M = 0$, some component of $M \cdot N$ joins a and b.*

11. THEOREM. *The necessary and sufficient condition for a locally connected compact continuum Z to be a unicoherent continuum is that Z is unicoherent about every simple closed curve J contained in Z.*

*Proof.** The condition is obviously necessary. To show that it is sufficient we assume that Z is not a unicoherent continuum. It is then the union of two continua whose divisor is not connected, and it is easy to show that it is the union of two *locally connected* continua M and N, such that $M \cdot N$ has a finite set of locally connected components, more than one in number.†

For some pair of these components, say A and B, $M - M \cdot N$ contains an open arc α whose end-points a and b are on A and B, respectively. Also $N - (A + B)$ contains an open arc β whose end-points c and d are on A and B, respectively. Now A contains a closed arc $\gamma = ac$ and B contains a closed arc $\delta = bd$. (γ or δ will be a point if $a = c$ or $b = d$, respectively.) Then $\alpha + \beta + \gamma + \delta$ is a simple closed curve J and Z is not unicoherent about J by § 4. This contradicts the hypothesis that Z is unicoherent about every simple closed curve which it contains, and so the theorem is proved.

12. In one of his articles ‡ Kuratowski collects in one theorem several characterizations of a locally connected compact unicoherent continuum. The following is the corresponding theorem for continua unicoherent about a simple closed curve. In the statement the word "arc" is always understood to include arcs which may lack one or both end-points and to include single points. The nature of the statements made is such that to any pair of arcs of J used, the complementary arcs are non-void. The proof has been omitted, since it is long and is merely a modification of Kuratowski's proof of the corresponding theorem.

THEOREM. *Let Z be a compact locally connected connected space containing the simple closed curve J. The following assertions are equivalent:*

---

* This theorem is also a consequence of § 3, Corollary 2, and a theorem of Borsuk (*loc. cit.*, p. 184), but this proof is retained on account of its brevity.

† See C. Kuratowski, "Sur quelques théorèmes fundamentaux de l'Analysis Situs," *Fundamenta Mathematicae*, Vol. 14, p. 307.

‡ C. Kuratowski, "Une caractérisation topologique de la surface de la sphère," *Fundamenta Mathematicae*, Vol. 13, p. 309.

(a) *Z is unicoherent about J.*

(b) *If M is a continuum containing in its interior an arc α of J, R is a component of $Z - M$ containing another arc β of J, and F is the frontier of R, some component of F joins the components λ and μ of $M - (\alpha + \beta)$.*

(c) *If A and B are disjoint closed sets containing the respective arcs α and β of J, there is a closed set C such that $C \cdot (A + B) = 0$, C is an $S(\alpha, \beta)$,* and some component of C joins the components of $J - (\alpha + \beta)$.*

(d) *If α and β are arcs of J and K is an irreducible closed $S(\alpha, \beta)$, some component of K joins the complementary arcs λ and μ of J.*

(e) *If α and β are arcs of J, λ and μ are their complementary arcs, A and B are disjoint closed sets such that $\alpha \cdot A \neq 0$, $\beta \cdot B \neq 0$, $\alpha \cdot B = \beta \cdot A = 0$, and $(A + B)(\lambda + \mu) = 0$, and neither A nor B is an $S(\lambda, \mu)$, then $A + B$ is not an $S(\lambda, \mu)$.*

13. Theorem. *Let Z be a locally connected compact metric space irreducible with respect to the property of being closed and unicoherent about the simple closed curve J. Let ab be a simple arc such that $(ab) \cdot J = a + b$ and let ab disconnect Z. Then ab divides Z into two components, each of which contains exactly one component of $J - (a + b)$.*

Proof. Let α and β be the components of $J - (a + b)$ and suppose that α lies in the component R of $Z - ab$. If $\beta \subset R$ and $M = R + ab$, $J \subset M$. The sum of ab and the remaining components of $Z - ab$ is a closed set N and, by § 7, since $M \cdot N$ is a simple arc, M is unicoherent about J, contrary to the hypothesis that Z is irreducibly unicoherent about J.

Suppose then that S is the component of $Z - ab$ containing β and set $N = S + ab$. Let P be the sum of ab and the components of $Z - ab$ other than R and S. Again $J \subset M + N$ and $P \cdot (M + N) = ab$; whence $M + N$ is unicoherent about J. This contradiction shows that there can be no third component and the theorem is proved.

14. Let R and S be the components of $Z - ab$ in the theorem of § 13. It is obvious that a and b are limiting points of both R and S. Let c be any other point of ab, d be a point of $J \cdot R$, and e be a point of $J \cdot S$. The points c, d, and e divide $J + ab$ into separated sets A and B; hence Z is the union of closed sets H and K such that $B \cdot H = A \cdot K = 0$. Since Z is unicoherent about J, $H \cdot K$ contains a continuum joining d and e and consequently containing c. Thus every point of ab is a limiting point of both R and S. Hence *no sub-set of ab disconnects Z.*

* I. e., a set met by every continuum joining α and β.

But Zippin * has shown that, if a locally connected compact continuum Z contains a simple closed curve J, every simple arc ab in Z such that $(ab) \cdot J = a + b$ is an irreducible cut of Z, and at least one such arc exists, then Z is the homeomorphic image of a two-dimensional simplex. Consequently we have this result:

THEOREM. *Let Z be a locally connected compact metric space which is irreducible with respect to the property of being closed and unicoherent about the simple closed curve J. Let every simple arc ab such that $(ab) \cdot J = a + b$ disconnect F. Then Z is the homeomorphic image of a two-dimensional simplex.*

This result can also be obtained by a direct proof by methods similar to those used by Whitney.

YALE UNIVERSITY.

* L. Zippin, "Characterization of the closed 2-cell," *Abstract No. 244, Bulletin of the American Mathematical Society*, Vol. 38 (1932), p. 803.

ON THE EXISTENCE OF TOTALLY IMPERFECT AND PUNCTIFORM CONNECTED SUBSETS IN A GIVEN CONTINUUM.

By Gordon T. Whyburn.

A set of points which contains no compact perfect subset is said to be totally imperfect and one containing no compact continuum is said to punctiform. It has been shown by F. Bernstein * that any euclidean space and by Hausdorff † that any separable, complete perfect space may be decomposed into two disjoint totally imperfect sets each having the power of the continuum. Sierpinski ‡ has shown that in any euclidean space E_n $(n > 1)$ the complement of every totally imperfect set (and indeed of every punctiform set) is connected, and thus every E_n $(n > 1)$ contains totally imperfect connected sets. Knaster and Kuratowski § later showed that even the Sierpinski triangle curve contains totally imperfect connected sets.

In this paper we shall obtain an extension of the above mentioned result of Sierpinski's which can be applied in arbitrary continua; and with its aid we are able to give necessary and sufficient conditions for the existence (1) of totally imperfect connected subsets containing an arbitrary point x in any locally compact continuum and (2) of punctiform connected subsets containing an arbitrary point x in any locally connected continuum. Also, we obtain necessary and sufficient conditions for the existence in hereditarily locally connected continua of punctiform connected subsets. Our result in this connection taken together with previously known results enables us to completely characterize those hereditarily locally connected continua ‖ in

* *Leipziger Berichte*, Vol. 60 (1908), p. 325.

† See *Mengenlehre* (1927), p. 176.

‡ *Fundamenta Mathematicae*, Vol. 1, p. 6, and Vol. 2, p. 94; see also Knaster and Kuratowski, *ibid.*, Vol. 2, p. 236.

§ *Bulletin of the American Mathematical Society*, Vol. 33 (1927), p. 106.

‖ We suppose all the point sets considered to be in a separable metric space. We use the notation $p \epsilon X$ to mean that p is a point of the point set X. The point p of a continuum M is a local separating point of M [See author's paper in *Monatshefte für Mathematik und Physik*, Vol. 36, (1929), p. 305] provided some neighborhood R of p exists such that $M \cdot \bar{R} - p$ is separated between some pair of points belonging to the component C of $M \cdot \bar{R}$ which contains p. If M is locally connected, then p will be a local separating point of M if and only if p is a cut point of some region (= connected open subset) in M. A continuum every subcontinuum of which is locally connected is said to be hereditarily locally connected.

which the 0-dimensional subsets, the totally disconnected subsets, and the punctiform subsets coincide, and thus to solve, in so far as hereditarily locally connected continua are concerned, a problem proposed by Menger.* Finally we shall give an example of a regular curve of order 3 which contains a punctiform connected subset.

2. We begin with a proposition comprising a strong generalization of the result of Sierpinski mentioned above.

THEOREM. *In any locally compact continuum M the complement of any totally imperfect set P plus the set L of all local separating points of M is connected.*

Proof. Set $M - P + L = E$ and suppose, contrary to the theorem, that $E = E_1 + E_2$, where E_1 and E_2 are mutually separated. Then $M - E$ $(= P - L \cdot P)$ contains a closed set K which separates a point p_1 of E_1 and a point p_2 of E_2 in M and such that if $M - K = N_1 + N_2$, then $\bar{N}_1 \cdot \bar{N}_2 \supset K$. Now since $K \subset P$ and P is totally imperfect, K is countable and hence K contains an isolated point x. But † x is a local separating point of M and hence belongs to L, contrary to $x \epsilon K \subset P - L \cdot P$.

COROLLARY 1. (Sierpinski) *In any euclidean space E_n $(n > 1)$, the complement of any totally imperfect set is connected.*

For no E_n $(n > 1)$ has any local separating point.

COROLLARY 2. *If M is locally connected, so also is E.*

For if $x \epsilon E$, let R be a region in M of diameter $< \epsilon$ about x. Then R is a locally compact continuum and $R \cdot P$ a totally imperfect subset of R and $R \cdot L$ is the set of local separating points of R. Hence $R - R \cdot P + R \cdot L = R \cdot E$ is connected.

COROLLARY 3. *Any locally compact continuum M is the sum of two connected sets having in common exactly the set L of local separating points of M.*

For by the result of Bernstein-Hausdorff, M is the sum of two disjoint totally imperfect sets M_1 and M_2. Thus $M_1 + L$ and $M_2 + L$ are connected, their sum is M and their common part is L.

---

* See K. Menger, *Kurventheorie*, Teubner, 1932, p. 370.

† See the author's paper in *Monatshefte für Mathematik und Physik* (1929), p. 308, Theorem 4.

COROLLARY 4. *Any locally compact continuum having no local separating points is the sum of two disjoint, connected and totally imperfect sets.*

3. THEOREM. *In order that the locally compact continuum M contain, for each point x of M, a totally imperfect connected set P_x containing x it is necessary and sufficient that the set L of local separating points of M be countable.*

To prove the sufficiency of the condition, set $M = M_1 + M_2$, where M_1 and M_2 are disjoint and totally imperfect, $M_1 \supset x$, and set $P_x = M_1 + L$. Then if L is countable, P_x is totally imperfect; and since $M - P_x \subset M_2$, our theorem in § 2 gives that P_x is connected.

The condition is also necessary. For if L is uncountable, then * there exist two points a and b of M, a subcontinuum N of M and a perfect subset P of $N \cdot L$ such that $N \cdot (\overline{M - N}) = a + b$ and every point of P separates a and b in N and is a point of order 2 of M. Since P is ordered † it contains ‡ a point x which is a limit point both of its predecessors and of its followers in the ordering of P. Then clearly every non-degenerate connected subset of M containing x contains either all points of P in some neighborhood of x which precede x or all points of P is some neighborhood of x which follow x; and in either case it contains a perfect subset of P. Thus if L is uncountable, not every point of M (indeed of L) can belong to a totally imperfect connected subset of M.

COROLLARY. *If M is locally connected and has only a countable number of local separating points, then each point of M belongs to some totally imperfect, connected and locally connected subset of M.*

Thus, in particular, the Sierpinski triangle curve (see Knaster and Kuratowski, *loc. cit.*) has the property stated in this corollary.

4. LEMMA. *If ab is an arc of local separating points of a locally connected continuum N, then either every inner point of ab separates a and b in N or ab contains a subarc st which is free § in some cyclic element * C of N.*

---

* See the author's paper in the *Transactions of the American Mathematical Society*, Vol. 32 (1930), pp. 444-454.

† *Loc. cit.*

‡ See Zarankiewicz, *Fundamenta Mathematicae*, Vol. 12 (1928), p. 119.

§ An arc st is said to be *free* in a continuum N provided that $st - (s + t)$ is an open subset of N. For definitions and properties of cyclic elements, see the author's paper "Concerning the structure of a continuous curve," *American Journal of Mathematics*, Vol. 50 (1928), pp. 167-194.

For if not every inner point of ab separates a and b in N, then there exists a cyclic element C of N containing a subarc xy of ab. Now at most a countable number of points of xy can be cut points of N, since $xy \subset C$; and any other point of xy must be a local separating point of C. Thus the non-local-separating points of C on xy are countable and hence * xy contains a subarc st which is free in C.

5. THEOREM. *In order that each point x of the locally compact and locally connected continuum M belong to some punctiform connected subset of M it is necessary and sufficient that the set L of local separating points of M be punctiform.*

The condition is sufficient. For let x be any point of M and let $M = M_1 + M_2$, where M_1 and M_2 are disjoint and totally imperfect and $M_1 \supset x$. Then if we set $P_x = M_1 + L$, it follows from § 2 that P_x is connected. But P_x is also punctiform. For suppose, on the contrary, that P_x contains a non-degenerate continuum N. Then we have $N = N \cdot M_1 + N \cdot L$; and since † L is an F_σ, therefore $N \cdot L$ is an F_σ and $N - N \cdot L$ is a G_δ. But since L is punctiform, $N - N \cdot L$ is dense in N and hence dense in itself. Thus by a well known theorem of Young's, $N - N \cdot L$ contains a perfect set, contrary to the fact that $N - N \cdot L \subseteq M_1$ and M_1 is totally imperfect.

To prove the necessity of the condition we suppose, on the contrary, that L contains some non-degenerate continuum K. Then since K is locally connected,‡ it contains an arc ab. Now if every inner point of ab separates a and b in M, let x be any inner point of ab which is a point of order 2 of M and let P be any non-degenerate connected subset of M containing x. Then clearly P contains at least one other point y of ab; but since every inner point of the subarc xy of ab must separate x and y in M we have $xy \subset P$, which proves that P is not punctiform. On the other hand, if not every inner point of ab separates a and b in M, then by § 4, ab contains a subarc st which is free in some cyclic element C of M. In this case it is clear that if x is any inner point of st which is not a cut point of M and P is any connected subset of M containing x, then since $P \cdot C$ is connected, $P \cdot C$ contains some subarc of st and hence P is not punctiform. Thus in either case M contains some point x which lies in no punctiform connected subset of M.

---

* See the author's paper in *Mathematische Annalen*, Vol. 102, p. 320, Cor. 1.

† *Loc. cit.*, *Theorem* 8, p. 318.

‡ This follows from the fact that all save possibly a countable number of the local separating points of any continuum M are points of order 2 of M; see my paper in *Monatshefte für Mathematik und Physik*, *loc. cit.*

6. Theorem. *In order that the hereditarily locally connected continuum H contain a punctiform connected subset it is necessary and sufficient that the set of all local separating points of some subcontinuum of H be punctiform.*

The sufficiency of the condition results immediately from § 5. For if the set of all local separating points of some subcontinuum N of H is punctiform, then N contains a punctiform connected set. To prove the necessity of the condition we suppose that H contains a punctiform connected set P and proceed to show that the set L of all local separating points of the subcontinuum $\bar{P} = N$ of H is punctiform. If this is not so, then L contains an arc ab; and since all save a countable number of points of L are points of order 2 of N, it is clear that the arc ab may be so chosen that a and b belong to P. Now since P is punctiform and connected, it cannot contain ab; and hence not every inner point of ab can separate a and b in N. Then by § 4, ab contains a subarc st which is free in some cyclic element C of N. But since $P \cdot C$ is connected and $\bar{P} \supset C$, it is seen at once that P must contain every point, save possibly one, of st, contrary to the fact that P is punctiform.

7. *Equivalent conditions.* The condition in the theorem just proved may be modified so as to take the following equivalent form:

(1) *In order that the hereditarily locally connected continuum H contain no punctiform connected subset it is necessary and sufficient that every cyclicly connected subcontinuum of H contain a free arc (of itself).*

For if H contains a cyclicly connected subcontinuum C which has no free arc, then by § 4, the set of local separating points of C is punctiform and hence, by § 5, C contains a punctiform connected set; and on the other hand, if H contains a punctiform connected set P, then the continuum $\bar{P}$ must have a non-degenerate cyclic element C, and just as in § 6 it follows that C can have no free arc.

Likewise, if we define a free-arc-continuum as a continuum in which the free arcs are everywhere dense, then by similar reasoning we can establish the following additional equivalent form:

(2) *In order that no connected subset of the hereditarily locally connected continuum H be punctiform it is necessary and sufficient that every cyclicly connected subcontinuum of H be a free-arc-continuum.*

8. It has been shown by the author * that every totally disconnected subset of any hereditarily locally connected continuum is 0-dimensional. This result combined with §§ 6 and 7 yields the following characterization of those

* See *American Journal of Mathematics*, Vol. 53 (1931), p. 379.

hereditarily locally connected continua whose 0-dimensional, totally disconnected, and punctiform subsets all coincide:

THEOREM. *In order that every punctiform subset of the hereditarily locally connected continuum H be 0-dimensional it is necessary and sufficient that every cyclicly connected subcontinuum of H have a free arc.*

9. EXAMPLE. *There exists a plane regular curve C of order 3 which has only a countable number of ramification points* (i. e., *points of order* ≥ 3) *and which contains a punctiform connected set.*

Let E be any plane continuum having the following property: (α) every maximal free arc A in E is contained in exactly one simple closed curve $J(A)$ in E such that $J(A) = A + B$, where B is also a free arc. Let us then define the set $T(E)$ to be the continuum obtained from E by taking each maximal free arc A in K and (i) subdivide it into a finite number of subarcs each of diameter $< \frac{1}{4}\delta(A)$, (ii) on each such subarc ab choose a non-dense perfect set P containing the points a and b, and (iii) for each maximal open interval xy complementary to P in ab, add on to E an arc xoy of diameter $< 2\delta(xy)$ which lies except for x and y wholly within $J(A)$ and in the complement of E, all the arcs xoy being so chosen that no two of them have any common points. Then clearly $T(E)$ will be a continuum likewise having property (α).

We now define the curve C as follows. Let K_0 be a unit circle, and let two complementary semicircular arcs on K_0 be designated as the free arcs in K_0. Set $T(K_0) = K_1$, $T(K_1) = K_2, \cdots, T(K_n) = K_{n+1}, \cdots$. Finally, let $K = \sum_1^\infty K_n$, and set $C = \bar{K}$. Then C has all the desired properties. That every point of K is a point of order ≤ 3 is immediately seen; and if x is any point of $\bar{K} - K$, ϵ is any positive number, and n is an integer such that $1/n < \epsilon$, then x is enclosed by some simple closed curve $J(A)$ in K_{n+1}, where A is a maximal free arc in K_n. And since $\delta[J(A)] < 1/n < \epsilon$ and $J(A) \cdot \overline{C - J(A)}$ consists of just the two end points of A, it follows that x is a point of order two of C. Thus C is a regular curve of order 3; and the ramification points of C are countable, because, for each n, the ramification points of C belonging to K_n are countable and it was just shown that no point of $C - K$ can be a ramification point. Finally, to see that C contains a punctiform connected set, in view of §§ 4 and 5 or § 7 above we have only to note that C is cyclicly connected and that since K_n contains no free arc of diameter $> 1/n$, C can contain no free arc at all.

It has previously been shown by the author * that no regular curve of order ≤ 3 can contain a totally imperfect connected subset. Thus we have the following situation: a curve of order ≤ 2 must be either an arc or a simple closed curve; a curve of order ≤ 3 can contain no totally imperfect connected subset but may contain (e. g., the curve C above) a punctiform connected subset; a curve of order 4 (e. g., the Sierpinski triangle curve) may contain totally imperfect connected subsets.

Incidentally the curve C just described yields negative answers to two questions previously proposed by the author.†

THE JOHNS HOPKINS UNIVERSITY.

* See *Bulletin of the American Mathematical Society*, Vol. 35 (1929), p. 223.

† *Loc. cit.*, p. 224; and *Fundamenta Mathematicae*, Vol. 12, p. 294.

PROBLEMS OF APPROXIMATION WITH INTEGRAL AUXILIARY CONDITIONS.*

By Dunham Jackson.

---

1. **Introduction.** The problem of approximating a given function by other functions of specified type may be varied by requiring that the approximating function shall satisfy auxiliary conditions of one form or another. Questions of this sort have been considered in various connections. In the present paper, auxiliary conditions will be imposed which require that certain definite integrals involving the approximating function agree exactly in value with the corresponding integrals in terms of the function to be approximated. Two problems of the type suggested will be dealt with in the next two sections, and special features of these problems will be further discussed in the concluding sections.

2. **Linear auxiliary conditions.** Let $f(x)$ be a given function, continuous for $a \leqq x \leqq b$. Let $\rho(x)$ and $\sigma(x)$ be non-negative functions summable over the interval (a, b), the latter being positive over a set of positive measure in every subinterval, and the former positive at least over some set of positive measure in (a, b); further restrictions will be imposed on $\rho(x)$ and $\sigma(x)$ as occasion arises. Let $\phi_1(x), \phi_2(x), \cdots, \phi_N(x)$ be N given functions defined for $a \leqq x \leqq b$, for simplicity continuous (though it will be apparent that this restriction can be relaxed), and linearly independent. A polynomial $P_n(x)$, of the n-th degree,† is to be defined as an approximating function by the requirement that the integral

$$\int_a^b \rho(x) \mid f(x) - P_n(x) \mid^m dx \tag{1}$$

shall be a minimum, subject to the condition that ‡

---

* Presented to the American Mathematical Society at Ames, Iowa, November 26, 1932.

† The words "of the n-th degree" will be understood throughout to mean "of the n-th degree at most."

‡ It would perhaps be most natural in first setting up the problem to think of $\rho(x)$ and $\sigma(x)$ as identical, or else to adopt an even more general formulation than the one here proposed; a middle ground has been chosen for the sake of realizing certain simplifications, and at the same time leaving open at least the alternatives $\sigma \equiv \rho$ and $\sigma \equiv 1$ and furthermore avoiding confusion of secondary hypotheses which may be necessary or convenient in the case of one weight function and irrelevant for the other.

$$(2) \quad \int_a^b \sigma(x)\phi_i(x)P_n(x)\,dx = \int_a^b \sigma(x)\phi_i(x)f(x)\,dx, \quad (i = 1, 2, \cdots, N),$$

the exponent m being a given positive number. It will be shown that such an approximating polynomial exists, at least for n sufficiently large, and that under appropriate hypotheses it converges uniformly toward $f(x)$ as n becomes infinite.

An essential preliminary is the following:

LEMMA I. *There exist polynomials* $\pi_1(x), \cdots, \pi_N(x)$ (*of degrees not specified*) *such that the determinant*

$$(3) \quad \begin{vmatrix} \int_a^b \sigma(x)\phi_1(x)\pi_1(x)\,dx & \cdots & \int_a^b \sigma(x)\phi_1(x)\pi_N(x)\,dx \\ \cdot & \cdots & \cdot \\ \int_a^b \sigma(x)\phi_N(x)\pi_1(x)\,dx & \cdots & \int_a^b \sigma(x)\phi_N(x)\pi_N(x)\,dx \end{vmatrix}$$

is different from zero.

In different words, if the other conditions remain as stated, vanishing of the determinant for every choice of the polynomials π_i would imply that the ϕ's are linearly dependent. In this form the assertion is a generalization of the fundamental fact, constituting the case $N = 1$, that a single continuous function orthogonal to every polynomial (with respect to a weight function $\sigma(x)$ of the character specified) is identically zero.

A proof, if not already familiar, may be given as follows:

In accordance with Weierstrass's theorem let a sequence of approximating polynomials $\pi_{in}(x)$ be constructed for each of the functions ϕ_i so that $\lim_{n\to\infty} \pi_{in}(x) = \phi_i(x)$, uniformly for $a \leqq x \leqq b$. If substitution of the polynomials $\pi_{1n}(x), \cdots, \pi_{Nn}(x)$ for $\pi_1(x), \cdots, \pi_N(x)$ makes the determinant zero for each value of n, it must be in the limit that

$$\begin{vmatrix} \int_a^b \sigma\phi_1^2\,dx & \cdots & \int_a^b \sigma\phi_1\phi_N\,dx \\ \cdot & \cdots & \cdot \\ \int_a^b \sigma\phi_N\phi_1\,dx & \cdots & \int_a^b \sigma\phi_N^2\,dx \end{vmatrix} = 0,$$

and the linear dependence of the ϕ's is thereby established.

(The fact that the vanishing of the last determinant is a sufficient condition for linear dependence is recognized as readily as in the special case $\sigma \equiv 1$: If the determinant is zero it is possible to find coefficients $c_1, \cdots, c_N$,

not all zero, to satisfy the simultaneous equations $\Sigma_j c_j \int \sigma\phi_i\phi_j = 0$, $(i = 1, 2, \cdots, N)$, i. e. $\int \sigma\phi_i\psi = 0$ if $\psi = \Sigma c_j\phi_j$; addition of these equations after multiplication by $c_1, \cdots, c_N$ respectively gives $\int \sigma\psi^2 = 0$, whence, as ψ is continuous and σ is positive on a set of positive measure in every subinterval, it follows that $\psi \equiv c_1\phi_1 + \cdots + c_N\phi_N \equiv 0$.)

The statement and proof of the lemma are obviously not restricted to polynomials, but apply equally well if the π's are understood to be linear combinations formed from any set of integrable functions in terms of which an arbitrary continuous function can be uniformly approximated. In particular, the lemma holds for an interval of length 2π if the ϕ's are continuous functions of period 2π and polynomials are replaced by trigonometric sums.

The lemma answers in the first place the question whether polynomials satisfying the auxiliary conditions (2) exist at all. For if polynomials $\pi_1(x), \cdots, \pi_N(x)$ are chosen so that the determinant (3) has a value different from zero, and if $\Pi(x)$ is any polynomial whatever, the system of equations

$$c_1 \int \sigma\phi_i\pi_1 + \cdots + c_N \int \sigma\phi_i\pi_N = \int \sigma\phi_i f - \int \sigma\phi_i\Pi, \quad (i = 1, 2, \cdots, N),$$

can be solved for the c's, and then $P \equiv \Pi + c_1\pi_1 + \cdots + c_N\pi_N$ is a polynomial having the desired property. If n_0 is the exponent of the highest power of x occurring in any of the π's, the coefficients of any higher powers of x in P are the same as in Π, and are completely arbitrary, and the auxiliary conditions are satisfied by infinitely many polynomials of any specified degree higher than n_0. (One would not expect to be able to satisfy the conditions in general by polynomials with fewer than N coefficients, and it is readily seen that in particular cases it may be necessary to resort to higher values of n_0 and n.)

It is possible then to answer the further question as to the existence of a minimizing polynomial of specified degree $n \geqq n_0$ for the integral (1), subject to the conditions (2). For the integral (1) is a continuous function of the coefficients in the polynomial; it is possible to mark off in the $(n+1)$-dimensional space of these coefficients a closed domain * within which the

* See, e. g. D. Jackson, "A generalized problem in weighted approximation," *Transactions of the American Mathematical Society*, Vol. 26 (1924), pp. 133-154, pp. 133-139; "Note on the convergence of a sequence of approximating polynomials," *Bulletin of the American Mathematical Society*, Vol. 37 (1931), pp. 69-72, p. 70.

coefficients of any polynomial bringing the value of the integral near its lower bound must be sought; and the conditions (2) define a closed subset of this domain, since if each polynomial of a sequence satisfies these equations and if the sequence uniformly approaches a limit the equations will be satisfied in the limit. Consequently there is at least one polynomial of the n-th degree satisfying (2) and minimizing (1). If $m > 1$ there is just one such polynomial, by the argument that is usual in similar cases,* the only point requiring special notice being the fact that if each of two polynomials satisfies (2) their average does likewise.

A step toward a convergence proof is represented by

LEMMA II. *The continuous function* $f(x)$ *being given, if there exist polynomials* $p_n(x)$, $(n = 0, 1, 2, \cdots)$, *each of degree indicated by its subscript* (i. e., as already noted, of that degree at most), *such that*

$$| f(x) - p_n(x) | \leqq \epsilon_n$$

for $a \leqq x \leqq b$, *there exist polynomials* $q_n(x)$ *of corresponding degree, for all values of* n *from a certain point on, satisfying the auxiliary conditions, and approximating* $f(x)$ *so that*

$$| f(x) - q_n(x) | \leqq K \epsilon_n, \tag{4}$$

where K *is independent of* n.

Let

$$\int_a^b \sigma(x)\phi_i(x)f(x)\,dx - \int_a^b \sigma(x)\phi_i(x)p_n(x)\,dx = h_{in}.$$

Let $\pi_1(x), \cdots, \pi_N(x)$ be chosen as before so that the determinant (3) is different from zero. For each n, let N coefficients $c_{n1}, \cdots, c_{nN}$ be defined by the system of equations

$$\sum_{j=1}^{N} c_{nj} \int_a^b \sigma(x)\phi_i(x)\pi_j(x)\,dx = h_{in}$$

with this non-vanishing determinant. The polynomial

$$q_n(x) = p_n(x) + c_{n1}\pi_1(x) + \cdots + c_{nN}\pi_N(x)$$

satisfies the conditions

$$\int_a^b \sigma(x)\phi_i(x)q_n(x)\,dx = \int_a^b \sigma(x)\phi_i(x)f(x)\,dx,$$

* See e. g. *Transactions of the American Mathematical Society, loc. cit.*, pp. 137-138; *Bulletin of the American Mathematical Society, loc. cit.*, p. 70.

and is of the n-th degree if $n \geqq n_0$, where n_0 still denotes the exponent of the highest power of x occurring in $\pi_1, \cdots, \pi_N$. If the c's are expressed by Cramer's rule each is a linear combination of the h's with coefficients which are independent of n, and as

$$| h_{in} | \leqq G\epsilon_n \int_a^b \sigma(x)\,dx$$

if G is a common upper bound for the functions $| \phi_i(x) |$, the assertion in the lemma is justified, since $| q_n(x) - p_n(x) |$ is seen to have a constant multiple of ϵ_n as an upper bound, and

$$| f(x) - q_n(x) | \leqq | f(x) - p_n(x) | + | p_n(x) - q_n(x) |.$$

Let $\int_a^b \rho(x)\,dx$ be denoted by W. From (4) it appears that

$$\int_a^b \rho(x) \, | f(x) - q_n(x) |^m \, dx \leqq WK^m\epsilon_n{}^m.$$

Let $P_n(x)$ be the particular polynomial of the n-th degree which minimizes the integral (1) subject to the conditions (2), or one such polynomial, if the determination is not unique, and let γ_n be the corresponding minimum value of (1). Then it is certain that

$$\gamma_n \leqq WK^m\epsilon_n{}^m. \tag{5}$$

To complete the proof of convergence of $P_n(x)$ toward $f(x)$ it will be advantageous, instead of giving details at full length here, to take over the substance of a corresponding proof which has been given elsewhere. The reasoning of pp. 96-97 of the writer's *Colloquium*,* though not summarized there in precisely these terms, may be regarded as constituting a proof of the following proposition:

LEMMA III. *If $\rho(x)$ is non-negative and summable over (a, b), and $\rho(x) \geqq v > 0$ for $\alpha_0 \leqq x \leqq \beta_0$, where v is constant and $a \leqq \alpha_0 < \beta_0 \leqq b$, if $f(x)$ is a continuous function for $a \leqq x \leqq b$, $P_n(x)$ an arbitrary polynomial of the n-th degree, and*

$$g_n = \int_a^b \rho(x) \, | f(x) - P_n(x) |^m \, dx,$$

and if there exists a polynomial $p_n(x)$, of the n-th degree, such that

$$| f(x) - p_n(x) | \leqq \epsilon_n$$

---

* "The theory of approximation," *American Mathematical Society Colloquium Publications*, Vol. 11 (New York, 1930); cited here and subsequently as *Colloquium*.

for $\alpha_0 \leqq x \leqq \beta_0$, *then, for* $\alpha_0 \leqq x \leqq \beta_0$,

$$| f(x) - P_n(x) | \leqq B_0(n^2 g_n)^{1/m} + 5\epsilon_n,$$

where $B_0 = 4 \cdot 32^{1/m} [(\beta_0 - \alpha_0)v]^{-1/m}$.

For the present application the precise value of B_0 is immaterial, the important thing being that it does not depend on n or on any other specification with regard to the polynomial $P_n(x)$. It is assumed in the text of the passage cited, and will be granted in the application here, that $| f(x) - p_n(x) | \leqq \epsilon_n$ for $a \leqq x \leqq b$, but this hypothesis is not actually used in the proof of the lemma outside the interval (α_0, β_0).

Let the polynomial $P_n(x)$ in the lemma be identified with the minimizing polynomial previously discussed, so that the integral g_n is that previously denoted by γ_n, and let $\rho(x)$ be subjected to the hypothesis of the lemma. Since it is possible by Weierstrass's theorem to construct polynomials $p_n(x)$, for the purposes of the lemma, so that $\lim_{n\to\infty} \epsilon_n = 0$, $P_n(x)$ will converge uniformly toward $f(x)$ for $\alpha_0 \leqq x \leqq \beta_0$ if $\lim_{n\to\infty} n^2\gamma_n = 0$, and so, by virtue of (5), if $\lim_{n\to\infty} n^{2/m}\epsilon_n = 0$; in applying Lemmas II and III it may be assumed for economy of notation that $p_n(x)$ and ϵ_n are the same in both cases. Conditions under which polynomials $p_n(x)$ can be constructed so as to make $\lim_{n\to\infty} n^{2/m}\epsilon_n = 0$ are given by known theorems on polynomial approximation.*
The conclusions for $m \geqq 2$ may be stated in

THEOREM I. *If* $\rho(x)$ *satisfies the hypothesis of Lemma* III, *the minimizing polynomial* $P_n(x)$ *will converge uniformly toward* $f(x)$ *for* $\alpha_0 \leqq x \leqq \beta_0$ *if* $m > 2$ *and* $f(x)$ *has throughout* (a, b) *a modulus of continuity* $\omega(\delta)$ *such that* $\lim_{\delta\to 0} \omega(\delta)/\delta^{2/m} = 0$, *or if* $m = 2$ *and* $f(x)$ *has a continuous derivative for* $a \leqq x \leqq b$.

The less simple statement for $0 < m < 2$ need not be explicitly formulated.

Further information with regard to convergence can be obtained by reference to another lemma, proved in substance in the *Colloquium*, though not formally stated there: †

LEMMA IV. *Let* $\rho(x)$ *be non-negative and summable over* (a, b), *and*

---

* See e. g. *Colloquium*, pp. 13-18. For the application cf. p. 98 of the *Colloquium*.

† *Colloquium*, pp. 98-101. By re-examination of the proof it would be possible to arrive at a statement corresponding even more closely to that of Lemma III, but the formulation given is adequate for the problem under consideration.

let $\rho(x) \geqq v > 0$ *for* $\alpha_0 \leqq x \leqq \beta_0$, *where* v *is constant and* $a \leqq \alpha_0 < \beta_0 \leqq b$. *Let* $f(x)$ *be a continuous function for* $a \leqq x \leqq b$. *Let two sequences of polynomials* $P_n(x)$, $p_n(x)$ *be defined for* $n = 1, 2, \cdots$, *each polynomial being of the degree indicated by its subscript (at most), but otherwise arbitrary. Let* ϵ_n *be an upper bound for* $| f(x) - p_n(x) |$ *in* (α_0, β_0):

$$| f(x) - p_n(x) | \leqq \epsilon_n$$

for $\alpha_0 \leqq x \leqq \beta_0$. *Let*

$$g_n = \int_a^b \rho(x) \, | f(x) - P_n(x) |^m \, dx.$$

Let α *and* β *be any two numbers such that* $\alpha_0 < \alpha < \beta < \beta_0$, *and let* η *be any positive number. Then, for* $\alpha \leqq x \leqq \beta$,

$$| f(x) - P_n(x) | \leqq A_k n^{(1/m)+\eta} g_n^{1/m} + 5\epsilon_n,$$

where A_k *is independent of* n.

The subscript k is merely an index which enters incidentally in the course of the proof, and is perpetuated here only for the sake of convenience of comparison with the passage referred to. To yield the result in the form stated here, the reasoning is to be modified superficially by omitting the assumption that $g_n \leqq A\gamma_n$ near the middle of p. 99 of the *Colloquium*, replacing ϵ_n by $g_n^{1/m}$ in the next to the last displayed formula on that page, so that it reads

$$\mu_{nk} \leqq A_k n^\sigma g_n^{1/m},$$

and making corresponding adjustments in the subsequent details.

By intermediate steps analogous to those which led from Lemma III to Theorem I it is possible to pass from Lemma IV to *

THEOREM II. *If* $\rho(x)$ *satisfies the hypothesis of Lemma* IV (*identical with the corresponding hypothesis of Lemma* III) *and if* $\alpha_0 < \alpha < \beta < \beta_0$, *the minimizing polynomial* $P_n(x)$ *will converge uniformly toward* $f(x)$ *for* $\alpha \leqq x \leqq \beta$ *if* $m > 1$ *and* $f(x)$ *has throughout* (a, b) *a modulus of continuity* $\omega(\delta)$ *such that, for some* $\eta > 0$, $\lim_{\delta \to 0} \omega(\delta)/\delta^{(1/m)+\eta} = 0$.

The less simple results for smaller values of m, in the present case $0 < m \leqq 1$, are again omitted from the formal statement.

An incidental consequence of Theorem II, without reference to uniformity of convergence, is the

COROLLARY.† *If* $\rho(x)$ *is non-negative and summable over* (a, b), *the*

* Cf. *Colloquium*, p. 101.

† The statement of the Corollary on p. 101 of the *Colloquium* should have been restricted to interior points of (a, b).

hypothesis with regard to $f(x)$ and the definition of $P_n(x)$ being as in Theorem II, *$P_n(x)$ will converge toward $f(x)$ at any interior point of (a, b) where $\rho(x)$ is continuous and different from zero.*

3. **Non-linear auxiliary condition.** An illustrative problem analogous to that of the preceding section, but different in some of the details of its working out, arises if the set of auxiliary conditions (2) is replaced by the single condition

$$\int_a^b \sigma(x)[P_n(x)]^2\,dx = \int_a^b \sigma(x)[f(x)]^2\,dx, \tag{6}$$

which is quadratic with respect to $P_n(x)$. It is assumed as before that $\sigma(x)$ is non-negative and summable over (a, b), and positive on a set of positive measure in every subinterval of (a, b).

There is no doubt this time as to the existence of polynomials satisfying the auxiliary condition. If any polynomial is given which is not identically zero the condition is satisfied by a suitable constant multiple of it. There will be infinitely many polynomials satisfying (6) for any specified $n \geqq 1$, and among them there will be at least one for which the integral (1) is a minimum, by the reasoning that was used in the earlier existence proof. On the other hand, the proof of uniqueness breaks down, since the average of two different polynomials satisfying (6) is *not* such a polynomial. Throughout the rest of this section it will be understood that $P_n(x)$ for each n is *a* polynomial of the n-th degree minimizing (1), subject to the condition (6), without further inquiry as to whether the determination is unique.

Lemma II of the preceding section can be adapted immediately as

LEMMA V. *If the hypothesis of Lemma* II *is satisfied, with $\lim_{n\to\infty} \epsilon_n = 0$, there exist polynomials $q_n(x)$ of the n-th degree, for n sufficiently large, satisfying the present auxiliary condition, and approximating $f(x)$ so that*

$$| f(x) - q_n(x) | \leqq K\epsilon_n,$$

where K is independent of n.

The trivial case $f(x) \equiv 0$ being ruled out, let

$$\int_a^b \sigma(x)\,dx = I > 0, \qquad \int_a^b \sigma(x)[f(x)]^2\,dx = D > 0,$$

and let

$$h_n = \int_a^b \sigma(x)[f(x)]^2\,dx - \int_a^b \sigma(x)[p_n(x)]^2\,dx,$$

where $p_n(x)$ is the polynomial given by the hypothesis. If M is the maximum

of $|f(x)|$, then $|p_n(x)| \leqq 2M$, at least for values of n from a certain point on, $|f(x) + p_n(x)| \leqq 3M$, and

$$(7) \qquad |f^2 - p_n^2| = |(f + p_n)(f - p_n)| \leqq 3M\epsilon_n, \qquad |h_n| \leqq 3MI\epsilon_n.$$

Let $d_n = [1 - (h_n/D)]^{-\frac{1}{2}} - 1$, and let $q_n(x) = (1 + d_n)p_n(x)$. It appears from (7) that $|d_n| \leqq k\epsilon_n$, where k is independent of n, and hence

$$|q_n(x) - p_n(x)| = |d_n p_n(x)| \leqq 2Mk\epsilon_n,$$

$$|f(x) - q_n(x)| \leqq |f(x) - p_n(x)| + |p_n(x) - q_n(x)| \leqq (2Mk + 1)\epsilon_n,$$

when n is sufficiently large, while

$$\int_a^b \sigma(x)[q_n(x)]^2\,dx = (1 + d_n)^2 \int_a^b \sigma(x)[p_n(x)]^2\,dx$$

$$= [1 - (h_n/D)]^{-1}(D - h_n) = D.$$

So the conclusion of the lemma is established, with $K = 2Mk + 1$.

Repetition of the later stages of the convergence proofs already indicated leads to

THEOREM III. *The assertions of Theorems* I *and* II *and the Corollary of the latter theorem hold for the approximating polynomials $P_n(x)$ which minimize the integral* (1) *subject to the auxiliary condition* (6).

4. Linear auxiliary conditions, special theorem on trigonometric approximation. Corresponding to the problems treated above there are analogous problems of trigonometric approximation. If it is assumed that the weight function in the integral to be minimized has a positive lower bound everywhere the trigonometric case is somewhat easier to deal with than the other, as the need for special considerations relative to the ends of an interval does not arise.

In the integrals (1) and (2) or (1) and (6) let the interval be that from $-\pi$ to π, let $f(x)$ and, in the case of (2), the ϕ's be continuous and of period 2π, and let $P_n(x)$ be replaced by a trigonometric sum $T_n(x)$ of the n-th order, so that the problem is that of minimizing the integral

$$(8) \qquad \int_{-\pi}^{\pi} \rho(x)\,|f(x) - T_n(x)|^m\,dx$$

subject to the conditions

$$(9) \quad \int_{-\pi}^{\pi} \sigma(x)\phi_i(x)T_n(x)\,dx = \int_{-\pi}^{\pi} \sigma(x)\phi_i(x)f(x)\,dx, \quad (i = 1, 2, \cdots, N),$$

or the single condition

$$(10) \qquad \int_{-\pi}^{\pi} \sigma(x)[T_n(x)]^2\,dx = \int_{-\pi}^{\pi} \sigma(x)[f(x)]^2\,dx.$$

Let the summable functions $\rho(x)$ and $\sigma(x)$, of period 2π, be everywhere non-negative, let $\sigma(x)$ be positive on a set of positive measure in every interval, and let $\rho(x)$ have the positive constant v as a lower bound for all values of x. For any specified n, sufficiently large, there is at least one $T_n(x)$ satisfying the auxiliary condition or conditions and reducing (8) to a minimum, and if $m > 1$, in the case of (9), the minimizing sum is uniquely determined.

In combination with a lemma corresponding to Lemma II or Lemma V and general theorems on degree of approximation by trigonometric sums * the following † serves as basis for a discussion of convergence:

LEMMA VI. *If $\rho(x)$ is of period 2π and summable over a period, and $\rho(x) \geqq v > 0$ everywhere, if $f(x)$ is a continuous function of period 2π, $T_n(x)$ an arbitrary trigonometric sum of the n-th order, and*

$$g_n = \int_{-\pi}^{\pi} \rho(x) \mid f(x) - T_n(x) \mid^m dx,$$

and if there exists a trigonometric sum $t_n(x)$, of the n-th order, such that

$$\mid f(x) - t_n(x) \mid \leqq \epsilon_n$$

for all values of x, then, for all values of x,

$$\mid f(x) - T_n(x) \mid \leqq 4(ng_n/v)^{1/m} + 5\epsilon_n.$$

The conclusion with regard to convergence may be recorded as follows for $m \geqq 1$, the case $0 < m < 1$ being left without formal statement:

THEOREMS Ia, IIIa. *The minimizing sums for the integral* (8) *with the auxiliary conditions* (9) *or the single auxiliary condition* (10) *will converge uniformly toward $f(x)$ for all values of x if $m > 1$ and $f(x)$ has everywhere a modulus of continuity $\omega(\delta)$ such that $\lim_{\delta \to 0} \omega(\delta)/\delta^{1/m} = 0$, or if $m = 1$ and $f(x)$ has everywhere a continuous derivative.*

So much is a fairly obvious parallel to the earlier discussion. The main purpose of the present section is to obtain a further result with regard to convergence in the trigonometric case with linear auxiliary conditions (consideration of the quadratic auxiliary condition being reserved for the next section) for the particular exponent $m = 2$, by the use of the following known proposition: ‡

LEMMA VII. *If $f(x)$ is an absolutely continuous function of period 2π,*

---

* See e. g. *Colloquium*, pp. 2-12.

† For a proof, without formulation of the result in these terms, see *Colloquium*, pp. 87-88; cf. *Colloquium*, p. 84, Theorem II a.

‡ See L. Tonelli, *Serie trigonometriche*, Bologna, 1928, p. 223; *Colloquium*, p. 56, Theorem VI.

and $S_n(x)$ the partial sum of its Fourier series through terms of the n-th order, and if δ_n is defined by the equation

$$\delta_n = \int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 \, dx,$$

then $\lim_{n\to\infty} n\delta_n = 0.$

The application is made through the medium of another lemma, to be proved here:

LEMMA VIII. *If $S_n(x)$ is the partial sum of the Fourier series for $f(x)$ through terms of order n, and*

$$\int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 \, dx = \delta_n,$$

and if $[\sigma(x)]^2$ as well as $\sigma(x)$ is summable, there exist trigonometric sums $v_n(x)$ of corresponding order for all values of n from a certain point on, satisfying the conditions (9) with $v_n(x)$ written in place of $T_n(x)$, and approximating $f(x)$ in the mean so that

$$\int_{-\pi}^{\pi} [f(x) - v_n(x)]^2 \, dx \leqq K\delta_n,$$

where K is independent of n.

(As far as Lemma VIII by itself is concerned the hypothesis of absolute continuity of $f(x)$ is irrelevant; it is sufficient in fact that $f(x)$ and its square be summable.)

In analogy with a notation previously used, let

$$\int_{-\pi}^{\pi} \sigma(x)\phi_i(x)[f(x) - S_n(x)] \, dx = h_{in}.$$

By Schwarz's inequality

$$\begin{aligned} h^2_{in} &\leqq \int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 \, dx \int_{-\pi}^{\pi} [\sigma(x)\phi_i(x)]^2 \, dx \\ &\leqq G^2\delta_n \int_{-\pi}^{\pi} [\sigma(x)]^2 \, dx, \end{aligned}$$

if G again denotes an upper bound for the absolute values of the ϕ's. Let the square root of the value of the last integral be denoted by J; then $|h_{in}| \leqq GJ\delta_n^{1/2}$.

An argument corresponding to that used in the proof of Lemma II demonstrates the existence of a trigonometric sum $u_n(x)$, whose coefficients depend on n through the h's, but whose order has a fixed upper bound as n increases and so is less than n for n sufficiently large, such that

$$\int_{-\pi}^{\pi} \sigma(x)\phi_i(x)u_n(x)\,dx = h_{in}, \qquad (i = 1, 2, \cdots, N),$$

and the manner of construction of $u_n(x)$ shows that its absolute value does not exceed a multiple of the greatest of the N numbers $|h_{1n}|, \cdots, |h_{Nn}|$ by a quantity independent of n. Hence

$$|u_n(x)| \leqq K_1 \delta_n^{1/2}$$

with a coefficient K_1 independent of n. Furthermore, if $S_n(x) + u_n(x)$ is denoted by $v_n(x)$,

$$\int_{-\pi}^{\pi} \sigma(x)\phi_i(x)v_n(x)\,dx = \int_{-\pi}^{\pi} \sigma(x)\phi_i(x)f(x)\,dx.$$

But if $F_1(x)$ and $F_2(x)$ are any two functions which are summable together with their squares over a specified interval,

$$\left[\int (F_1 + F_2)^2\right]^{1/2} \leqq \left[\int F_1^2\right]^{1/2} + \left[\int F_2^2\right]^{1/2}$$

when the integrals are extended over the interval in question. and in the present instance

$$\begin{aligned} \left[\int_{-\pi}^{\pi}\{f(x) - v_n(x)\}^2\,dx\right]^{1/2} &= \left[\int_{-\pi}^{\pi}\{[f(x) - S_n(x)] + [-u_n(x)]\}^2\,dx\right]^{1/2} \\ (11) \qquad &\leqq \left[\int_{-\pi}^{\pi}\{f(x) - S_n(x)\}^2\,dx\right]^{1/2} + \left[\int_{-\pi}^{\pi}\{u_n(x)\}^2\,dx\right]^{1/2} \\ &\leqq \delta_n^{1/2} + (2\pi K_1^2\delta_n)^{1/2}, \end{aligned}$$

$$\int_{-\pi}^{\pi}[f(x) - v_n(x)]^2\,dx \leqq K\delta_n$$

with $K = [1 + (2\pi)^{1/2} K_1]^2$, whereby the truth of the lemma is established.

Let it be assumed now that $\rho(x)$ is bounded above, in addition to having a positive lower bound:

$$0 < v \leqq \rho(x) \leqq V$$

for all values of x. Then

$$\int_{-\pi}^{\pi} \rho(x)[f(x) - v_n(x)]^2\,dx \leqq VK\delta_n,$$

and if $T_n(x)$ is the trigonometric sum of the n-th order which minimizes (8) for $m = 2$, subject to the auxiliary conditions (9),

$$\gamma_n = \int_{-\pi}^{\pi} \rho(x)[f(x) - T_n(x)]^2\,dx \leqq VK\delta_n.$$

Combination of this inequality with Lemmas VI and VII yields at once

THEOREM IV. *If $\rho(x)$ is a bounded measurable function with a positive*

lower bound, and $\sigma(x)$ non-negative everywhere, positive on a set of positive measure in every interval, and summable with its square from $-\pi$ to π, the sum $T_n(x)$ which minimizes (8) for $m=2$, with the auxiliary conditions (9), will converge uniformly toward $f(x)$ for all values of x as n becomes infinite, if $f(x)$ is absolutely continuous.

5. **Non-linear auxiliary condition, special theorem on trigonometric approximation.** Finally, the conclusion of Theorem IV is to be extended to the trigonometric sums $T_n(x)$ which minimize (8), for $m=2$, subject to the quadratic auxiliary condition (10). There is no assertion now that the approximating sum is unique, and it is to be understood that $T_n(x)$ for each n is *a* sum which reduces (8) to the smallest value consistent with (10).

The new fact that is needed in preparation for the convergence proof is expressed in

LEMMA IX. *In the statement of Lemma* VIII *the linear conditions* (9) *may be replaced by the non-linear condition* (10), *if it is assumed, in addition to the hypotheses previously imposed, that $\sigma(x)$ is bounded.*

It may be supposed for simplicity, and without sacrifice of generality in the application for which the lemma is desired, that $f(x)$ is continuous. Then it is not only well known, but is an immediate consequence of the least-square property of the Fourier series together with Weierstrass's theorem on uniform approximation by trigonometric sums, that $\lim_{n\to\infty} \delta_n = 0$.

It follows at once from the definition of the Fourier coefficients that $f(x) - S_n(x)$ is orthogonal to each of the functions $1, \cos x, \cdots, \cos nx, \sin x, \cdots, \sin nx$, and so to $S_n(x)$ itself:

$$\int_{-\pi}^{\pi} [f(x) - S_n(x)]\, S_n(x)\,dx = 0, \qquad \int_{-\pi}^{\pi} f(x) S_n(x)\,dx = \int_{-\pi}^{\pi} [S_n(x)]^2\,dx.$$

A familiar corollary is that

$$\int_{-\pi}^{\pi} [f(x) - S_n(x)]^2\,dx = \int_{-\pi}^{\pi} [f(x)]^2\,dx - \int_{-\pi}^{\pi} [S_n(x)]^2\,dx,$$

whence, as the left-hand member is non-negative,

$$\int_{-\pi}^{\pi} [S_n(x)]^2\,dx \leqq \int_{-\pi}^{\pi} [f(x)]^2\,dx,$$

the last integral being independent of n. Let its value be denoted by D_1. Then

$$\begin{aligned}\int_{-\pi}^{\pi} [f(x) + S_n(x)]^2\,dx &= \int_{-\pi}^{\pi} f^2 dx + 2\int_{-\pi}^{\pi} f S_n dx + \int_{-\pi}^{\pi} S_n^2 dx \\ &= \int_{-\pi}^{\pi} f^2 dx + 3\int_{-\pi}^{\pi} S_n^2 dx \leqq 4D_1.\end{aligned}$$

By Schwarz's inequality,

$$[\int_{-\pi}^{\pi} | f^2 - S_n^2 | \, dx]^2 = [\int_{-\pi}^{\pi} | f - S_n | \; | f + S_n | \, dx]^2$$
$$\leqq \int_{-\pi}^{\pi} (f - S_n)^2 dx \int_{-\pi}^{\pi} (f + S_n)^2 dx \leqq 4D_1\delta_n,$$
$$\int_{-\pi}^{\pi} | f^2 - S_n^2 | \, dx \leqq 2D_1^{1/2}\delta_n^{1/2}.$$

Let

$$h_n = \int_{-\pi}^{\pi} \sigma(x) [f(x)]^2 \, dx - \int_{-\pi}^{\pi} \sigma(x) [S_n(x)]^2 \, dx,$$

and let $\sigma(x)$, now assumed to be bounded, have V_1 for an upper bound. Then

$$| h_n | \leqq 2V_1 D_1^{1/2} \delta_n^{1/2}.$$

In analogy with the concluding stages of the proof of Lemma V, let

$$d_n = [1 - (h_n/D)]^{-1/2} - 1,$$

where

$$D = \int_{-\pi}^{\pi} \sigma(x) [f(x)]^2 \, dx,$$

and let $v_n(x) = (1 + d_n) S_n(x)$. It is seen that $| d_n | \leqq k\delta_n^{1/2}$, with a coefficient k independent of n, and that

$$\int_{-\pi}^{\pi} \sigma(x) [v_n(x)]^2 \, dx = D.$$

Furthermore, by adaptation of (11),

$$[\int_{-\pi}^{\pi} \{f(x) - v_n(x)\}^2 dx]^{1/2} = [\int_{-\pi}^{\pi} \{[f(x) - S_n(x)] + [- d_n S_n(x)]\}^2 dx]^{1/2}$$
$$\leqq [\int_{-\pi}^{\pi} \{f(x) - S_n(x)\}^2 dx]^{1/2} + [\int_{-\pi}^{\pi} \{d_n S_n(x)\}^2 dx]^{1/2}$$
$$\leqq \delta_n^{1/2} + | d_n | D_1^{1/2} \leqq (1 + kD_1^{1/2}) \delta_n^{1/2}.$$

With $K = (1 + kD_1^{1/2})^2$, the statement of the lemma is justified.

By the minimizing property of $T_n(x)$, if $\rho(x)$ has the upper bound V,

$$\int_{-\pi}^{\pi} \rho(x) [f(x) - T_n(x)]^2 \, dx \leqq \int_{-\pi}^{\pi} \rho(x) [f(x) - v_n(x)]^2 \, dx \leqq VK\delta_n,$$

and application of Lemmas VI and VII now leads to

THEOREM V. *In the statement of Theorem* IV *the linear conditions* (9) *may be replaced by the non-linear condition* (10) *if it is assumed, in addition to the hypotheses previously imposed, that* $\sigma(x)$ *is bounded.*

THE UNIVERSITY OF MINNESOTA,
MINNEAPOLIS.

MATRICES CONJUGATE TO A GIVEN MATRIX WITH RESPECT TO ITS MINIMUM EQUATION.

By Elizabeth S. Sokolnikoff.

1. *Introduction.* A set of matrices "conjugate" to a given matrix was first defined by Taber in an attempt to generalize the conjugate of a quaternion. In his two long articles on matrices * Taber defines and discusses such conjugates for the third order matrix whose characteristic equation has distinct roots. He remarks that the discussion can be extended to give the definition of the conjugates of any matrix whose characteristic equation has distinct roots. It is noted also that his definition can be applied if the minimum equation of the matrix has distinct roots. Moreover, Taber shows that, if a quaternion be interpreted as a second order matrix, the conjugate by this new definition is precisely the quaternion conjugate as ordinarily defined. In 1921 Bennett † pointed out certain analogies existing between the conjugates of a matrix, as defined by Taber, and the conjugates of an algebraic number.

A definition for the conjugate matrices of a general matrix was first given by Franklin.‡ He defines the conjugate matrices of a given n-rowed matrix as any set of $n-1$ matrices which possesses the two properties: (1) the matrices are commutative as to multiplication, (2) the elementary symmetric functions of the given matrix and the set of $n-1$ matrices are equal to the elementary symmetric functions of the roots of the characteristic equation, when these latter functions are considered as scalar matrices. The fact that each matrix of the set will satisfy the characteristic equation follows from (1) and (2). Franklin exhibited one such set of matrices. If the matrix satisfies Taber's requirements, Franklin's definition coincides with Taber's.

The conjugate matrices discussed in this paper are defined with respect to the minimum equation of the matrix. The minimum equation possesses certain advantages over the characteristic equation. Although the minimum

* *American Journal of Mathematics*, Vol. 12 (1890), p. 337; Vol. 13 (1891), p. 157.

† *Annals of Mathematics*, Vol. 23 (1921), p. 91.

‡ *Annals of Mathematics*, Vol. 23 (1921), p. 97.

equation is not ordinarily irreducible in the field of the elements of the matrix, it is the equation of lowest degree satisfied by the matrix and is, therefore, analogous to the defining equation of an algebraic number. Moreover, the conjugates with respect to the minimum equation form, with the given matrix, a set of m roots of the equation of lowest degree, m, satisfied by the matrix. The minimum equations of the conjugate matrices are divisors of the minimum equation of the given matrix. This property is not possessed by the conjugates if the characteristic equation replaces the minimum equation in the discussion. A further strengthening of the analogy with the algebraic number conjugates is furnished by the property that the conjugates defined in this paper can be expressed as polynomials, with scalar coefficients, in the given matrix. Franklin's conjugates do not admit of such representation, although it is possible to define the conjugates with respect to the characteristic equation so that they have this property.

The results given in this paper hold for any complex number field, and, unless specifically stated otherwise, the discussion will be considered to apply to any complex number field which contains the elements of the matrices considered.

2. *Notation.* Because of the complicated expression of the canonical form of the general matrix, it will be convenient to use certain notations. Let I represent the identity matrix and let J represent the square matrix having zeros for all its elements except for ones down the first diagonal above the principal diagonal. Then, for any value of n less than the number of rows in J, J^n will represent a matrix having zeros for all its elements except for ones down the n-th diagonal above the principal diagonal. If J has p rows, it is evident that $J^n = 0$ for $n \geqq p$. The matrices in which we are interested will have $p_i \times p_i$ blocks of the form $A_i(p_i, s_i)$, defined by

$$(1) \qquad A_i(p_i, s_i) \equiv (a_{i0}I + a_{i1}J + a_{i2}J^2 + \cdots + a_{is_i}J^{s_i})_{p_i}$$
$$\equiv (a_{i0}I + \sum_{k=1}^{s_i} a_{ik}J^k)_{p_i}.$$

$A_i(1, 0)$ will be represented by $(a_{i0}I)$. We further define the matrix

$$(2) \quad A \equiv [A_1(p_1, s_1),\ A_2(p_2, s_2), \cdots,\ A_n(p_n, s_n)]$$
$$\equiv [(a_{10}I + \sum_{k=1}^{s_1} a_{1k}J^k)_{p_1},\ (a_{20}I + \sum_{k=1}^{s_2} a_{2k}J^k)_{p_2}, \cdots,\ (a_{n0}I + \sum_{k=1}^{s_n} a_{nk}J^k)_{p_n}]$$

as the matrix which has the blocks $A_i(p_i, s_i)$ down its principal diagonal, and has zeros for all elements not included in these blocks.

Let A and B be two matrices of the form (2) where corresponding blocks are of the same number of rows. Then

$$A \equiv [A_1(p_1, s_1),\ A_2(p_2, s_2), \cdots,\ A_n(p_n, s_n)]$$

and

$$B \equiv [B_1(p_1, t_1),\ B_2(p_2, t_2), \cdots,\ B_n(p_n, t_n)].$$

It is clear that $A + B \equiv [A_1 + B_1,\ A_2 + B_2, \cdots;\ A_n + B_n]$, where

$$A_i + B_i = (\{a_{i0} + b_{i0}\}I + \sum_{k=1}^{p_i-1} (a_{ik} + b_{ik})J^k)_{p_i} \tag{3}$$

and $a_{ik} = 0$ if $k \geq s_i$ and $b_{ik} = 0$ if $k \geq t_i$. Similarly,

$$AB \equiv [A_1B_1,\ A_2B_2, \cdots,\ A_nB_n],$$

in which

$$A_iB_i = (a_{i0}b_{i0}I + \sum_{k=1}^{p_i-1} c_{ik}J^k)_{p_i}. \tag{4}$$

The c_{ik} are given by the expressions $c_{ik} = \sum\limits_{u+v=k} a_{iu}b_{iv} = \sum\limits_{u+v=k} b_{iv}a_{iu}$. From these last relations it appears that, when the corresponding blocks have the same size, the matrices are commutative as to multiplication.

The structure of $A + B$ and AB indicates that the discussion of sums and products of matrices of this type can often be restricted to the investigation of the sums and products of typical blocks. The products arising in our discussions will be products of two or more blocks of the type $A_i(p_i, 1)$ or $A_i(p_i, 0)$.

Consider the product of m of these $p_i \times p_i$ blocks $A_i^{(k)}$, where $a_1^{(k)}$ may be zero. If the subscripts i and p_i are omitted, this product has the form

$$\prod_{k=1}^{m} (a_0^{(k)}I + a_1^{(k)}J) \equiv (b_0I + \sum_{i=1}^{m} b_iJ^i). \tag{5}$$

The b_j are symmetric functions of the $a^{(k)}$, given by $b_0 = \prod\limits_{k=1}^{m} a_0^{(k)}$ and

$$b_j = \sum a_1^{(1)}a_1^{(2)} \cdots a_1^{(j)}a_0^{(j+1)}a_0^{(j+2)} \cdots a_0^{(m)}, \qquad (j = 1, 2, \cdots, m).$$

In particular,

$$(a_0I + a_1J)^m = (a_0^mI + \sum_{k=1}^{m} \binom{m}{k} a_0^{m-k}a_1^kJ^k). \tag{6}$$

It is evident that symmetric functions of m such blocks, $(a_0^{(k)}I + a_1^kJ)$, will be blocks whose elements are symmetric functions of the $a^{(k)}$.

3. *Definition of matrices conjugate to a given matrix with respect to its minimum equation.* Let M be a matrix whose minimum equation, $g(x) = 0$, is of degree m and has the q distinct roots $\rho_1, \rho_2, \cdots, \rho_q$ of respective multiplicities $\pi_1, \pi_2, \cdots, \pi_q$. A set of $m - 1$ matrices $M_1, M_2, \cdots, M_{m-1}$ will be

called a set of matrices conjugate to M with respect to its minimum equation if the set has the properties:

(i) Each M_i is expressible as a polynomial in M with coefficients in the field formed by adjoining the roots ρ_i and the π_i-th roots of unity to the field of the elements of M.

(ii) The elementary symmetric functions of $M, M_1, M_2, \cdots, M_{m-1}$ coincide with the elementary symmetric functions of the roots of $g(x) = 0$, when these latter functions are considered as scalar matrices.

From (i) it is obvious that the M_i are commutative as to multiplication. Moreover, by the use of (ii) and this commutative property, it can be shown that each M_i satisfies $g(x) = 0$.

There will not be a unique set of $m - 1$ matrices which has the properties (i) and (ii). We shall exhibit and discuss one set which possesses these properties. The structure of matrices forming other sets will be discussed in a later section. Unless another set is specifically described, the conjugate matrices discussed will be the set of matrices whose structure is described in this section.

For convenience, the matrix M will be used in the form $M = PR_0P^{-1}$, where

$$(7) \qquad R_0 = [(r_1I + J)_{p_1}, (r_2I + J)_{p_2}, \cdots, (r_aI + J)_{p_a}, (r_{a+1}I), \cdots, (r_bI)].$$

The r_i may not be distinct, but the distinct numbers among them will be the ρ_j. If it is assumed that $\pi_1 \geq \pi_2 \geq \cdots \geq \pi_q$ and $p_1 \geq p_2 \geq \cdots \geq p_a$, then $r_1 = \rho_1$, $r_2 = \rho_1$ or ρ_2, etc.

The set of $m - 1$ matrices conjugate to M with respect to its minimum equation will be defined as the set of matrices $M_i = PR_iP^{-1}$, $(i = 1, 2, \cdots, m - 1)$, in which the R_i are obtained from R_0 by the method described below. It will be assumed that all subscripts on ρ and π have been reduced modulo q to the set $1, 2, \cdots, q$.

(a) Consider any block of the form $(r_kI + J)_{p_k}$. As a root of $g(x) = 0$, $r_k \equiv \rho_j$ has multiplicity $\pi_j \geq p_k$. In order to form R_1, replace the block $(\rho_jI + J)_{p_k}$ by $(\rho_jI + \omega_{\pi_j}J)_{p_k}$, where ω_{π_j} is a primitive π_j-th root of unity. In general, in order to form R_i $(i = 1, 2, \cdots, \pi_j - 1)$, replace $(\rho_jI + J)_{p_k}$ by $(\rho_jI + \omega_{\pi_j}{}^iJ)_{p_k}$. The next π_{j+1} conjugates are formed by replacing the block $(\rho_jI + J)_{p_k}$ by $(\rho_{j+1}I)_{p_k}$, the following π_{j+2} conjugates by replacing the block $(\rho_jI + J)_{p_k}$ by $(\rho_{j+2}I)_{p_k}$, etc. The final π_{j+q-1} conjugates are formed by replacing the block $(\rho_jI + J)_{p_k}$ by $(\rho_{j+q-1}I)_{p_k}$.

(b) Consider any single element block (r_lI). As a root of $g(x) = 0$

$r_l \equiv \rho_j$ has multiplicity π_j. In the first $\pi_j - 1$ conjugates leave the block unaltered. In the succeeding π_{j+1} conjugates replace it by $(\rho_{j+1}I)$, in the following π_{j+2} conjugates replace it by $(\rho_{j+2}I)$, etc.

For the special case in which the roots of $g(x) = 0$ are all distinct, $q = m$ and $\pi_1 = \pi_2 = \cdots = \pi_m = 1$. In this case the R_i are given by $R_i = [(\rho_{i+1}I)_{p_1}, (\rho_{i+2}I)_{p_2}, \cdots, (\rho_i I)_{p_m}]$, $(i = 1, 2, \cdots, m-1)$. It is evident that each conjugate matrix has $g(x) = 0$ for its minimum equation. In general, as will be proved later, the minimum equation of any conjugate matrix is a divisor of $g(x) = 0$.

The following example is given to exhibit the form of the conjugate matrices of a particular matrix. Let P be any non-singular square matrix of order 8, and let $R_0 = [(I + J)_3, (I + J)_2, (-I + J)_2, (2I)]$. Then the minimum equation of the matrix $M = PR_0P^{-1}$ is $(x-1)^3(x+1)^2(x-2) = 0$. The conjugate matrices are $M_i = PR_iP^{-1}$, where the R_i are given by

$$R_1 = [(I + \omega_1 J)_3, (I + \omega_1 J)_2, (-I - J)_2, (I)];$$
$$R_2 = [(I + \omega_2 J)_3, (I + \omega_2 J)_2, (2I)_2, (I)]; \; R_3 = [(-I)_3, (-I)_2, (I)_2, (I)];$$
$$R_4 = [(-I)_3, (-I)_2, (I)_2, (-I)]; \; R_5 = [(2I)_3, (2I)_2, (I)_2, (-I)].$$

4. *Properties of the conjugate matrices.* It may be noted that $M_iM_j = PR_iP^{-1}PR_jP^{-1} = PR_iR_jP^{-1}$, and that $M_i + M_j = P\{R_i + R_j\}P^{-1}$. It follows that, if $\theta(M)$ is a polynomial in M, then $\theta(M_i) = P\theta(R_i)P^{-1}$. In view of this relation it will be sufficient to prove that the set of matrices R_0, R_1, $R_2, \cdots, R_{m-1}$ possesses the properties (i) and (ii).

THEOREM 1. *Let R_0 be a matrix whose minimum equation, of degree m, has the distinct roots $\rho_1, \rho_2, \cdots, \rho_q$ of respective multiplicities $\pi_1, \pi_2, \cdots, \pi_q$. The conjugate matrices R_i, $(i = 1, 2, \cdots, m-1)$, can be expressed as polynomials in R_0 of the form*

$$R_i = \theta_i(R_0) \equiv \xi_0^{(i)} I + \xi_1^{(i)} R_0 + \cdots + \xi_{m-1}^{(i)} R_0^{m-1}, \tag{8}$$

where the $\xi_j^{(i)}$ are in the field formed by the adjunction to the field F (the field of the elements of M) of the ρ_k and the π_k-th roots of unity, $(k = 1, 2, \cdots, q)$.

By definition

$$R_0 = [(r_1 I + J)_{p_1}, (r_2 I + J)_{p_2}, \cdots, (r_a I + J)_{p_a}, (r_{a+1} I), \cdots, (r_b I)], \tag{9}$$

and

$$R_i = [(s_1 I + t_1 J)_{p_1}, (s_2 I + t_2 J)_{p_2}, \cdots, (s_a I + t_a J)_{p_a}, (s_{a+1} I), \cdots, (s_b I)], \tag{10}$$

in which the s_j are among the ρ_k and the t_j are π_k-th roots of unity or zero. Since the blocks of R_0 are of the form (1) with s_i equal to 1 or 0, the powers of R_0 will have blocks of the form (6). The conditions on the $\xi_j^{(i)}$ will be obtained by equating corresponding elements of the matrix R_i and the matrix $\theta_i(R_0)$. The conditions are

$$(11)\qquad \left\{\begin{aligned} &\sum_{j=0}^{m-1} \xi_j^{(i)} \rho_k^{j} = s_l \\ &\sum_{j=1}^{m-1} \xi_j^{(i)} \binom{j}{1} \rho_k^{j-1} = t_l \\ &\sum_{j=n}^{m-1} \xi_j^{(i)} \binom{j}{n} \rho_k^{j-n} = 0, \quad (n = 2, 3, \cdots, \pi_k - 1), \end{aligned}\right.$$

where $k = 1, 2, \cdots, q$. Since $m = \pi_1 + \pi_2 + \cdots + \pi_q$, (11) gives a set of m linear non-homogeneous equations in the $\xi_j^{(i)}$. The matrix of the coefficients of these equations will have the following structure. Let G_k represent the $l \times m$ matrix which has unity for the element in the k-th column and zeros for all other elements. The form for the general, j-th, row of the first π_1 rows is given by $\sum_{k=1}^{m} g_{jk}\rho_1^{k-j} G_k$, $(j = 1, 2, \cdots, \pi_1)$, where $g_{jk} = 1$ if $j = 1$, $g_{jk} = 0$ if $k < j$, and $g_{jk} = \binom{k-1}{j-1}$ for all other values of k and j. The following π_2 rows are of the same form except that ρ_2 replaces ρ_1 and j takes the values $1, 2, \cdots, \pi_2$. The remaining $\pi_3, \pi_4, \cdots, \pi_q$ rows are of the same type with ρ_1 replaced successively by $\rho_3, \rho_4, \cdots, \rho_q$. The determinant Δ of this matrix has the value

$$(12)\qquad \Delta = \prod_{i>j}^{q} (\rho_i - \rho_j)^{\pi_i \pi_j}.$$

Since the ρ_k are distinct, the relation (12) shows that $\Delta \neq 0$. It follows that the equations (11) can be solved for the $\xi_j^{(i)}$.

It is of interest to note that, since R_i contains elements which are the π_h-th roots of unity only for $i < \pi_h$, the $\xi_j^{(i)}$ for $i \geq \pi_h$ will be in the field formed by the adjunction to F of the ρ_k, $(k = 1, 2, \cdots, q)$ and the π_l-th roots of unity $(l = 1, 2, \cdots, h - 1)$. In the special case in which $\pi_1 = \pi_2 = \cdots = \pi_q = 1$, the $\xi_j^{(i)}$ are, for all values of i, in the field formed by adjoining the ρ_k to F. In this case the determinant Δ reduces to the Vandermondian $| r_j^{i-1} |$, $(i, j = 1, 2, \cdots, m)$.

In the particular case in which the minimum equation is a quadratic equation with coefficients in the field of the elements of M, the coefficients of the polynomial expression for the conjugate matrix M_1 will be in this field, even if the roots of the minimum equation are not in the field. For, if

$g(x) \equiv x^2 - a_1x + a_0 = 0$, it can be shown easily that $M_1 = a_1I - M$. That the $\xi_j^{(i)}$ do not always lie in the field of the elements of M can be shown by a simple example. Let $M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & 7 & -3 \end{pmatrix}$; then $R_0 = [(I), (2^{1/2}I), (-2^{1/2}I)]$ and $R_1 = [(2^{1/2}I), (-2^{1/2}I), (I)]$ and $R_2 = [(-2^{1/2}I), (I), (2^{1/2}I)]$. The polynomial expressions for the conjugates are

$$M_1 = \frac{1 + 6(2)^{1/2}}{2} I - \frac{2 + 2^{1/2}}{4} M - \frac{7(2)^{1/2}}{4} M^2$$

and

$$M_2 = \frac{1 - 6(2)^{1/2}}{2} I - \frac{2 - 2^{1/2}}{4} M + \frac{7(2)^{1/2}}{4} M^2.$$

THEOREM 2. *The elementary symmetric functions E_i, $(i = 1, 2, \cdots, m)$, of the matrices $R_0, R_1, \cdots, R_{m-1}$ have the form $E_i = e_iI$, where the e_i are the elementary symmetric functions of the roots of $g(x) = 0$.*

Since the functions E_i are sums of products of matrices of the form discussed in § 2, it will be sufficient to consider a typical block and its conjugate blocks. Let us first consider a block of the form $(r_iI + J)_p$, having the conjugate blocks $(r_iI + \omega^jJ)_p$, $(j = 1, 2, \cdots, \pi - 1)$, and $(r_kI)_p$, $(k = 1, 2, \cdots, m - \pi)$, where $r_k \neq r_i$. π is the multiplicity of r_i considered as a root of $g(x) = 0$ and therefore $\pi \geq p$. The n-th elementary symmetric function of these blocks will have the form $(e_nI + \sum_{i=1}^{n} c_{ni}J^i)_p$. The c_{ni} are homogeneous symmetric functions of the π-th roots of unity. Moreover, the c_{ni} are all of degree less than p, and, by use of the fundamental theorem on symmetric functions, each c_{ni} can be expressed as a polynomial without a constant term in the first $p - 1$ elementary symmetric functions of the π-th roots of unity. Since $\pi \geq p$, these functions c_{ni} are all zero. It follows that the block representing the n-th elementary symmetric function of the typical block $(r_iI + J)_p$ and its conjugates, has the form $(e_nI)_p$. Obviously the n-th elementary symmetric function of any block (r_iI) and its conjugate blocks will have the same form. It follows that $E_i = e_iI$.

COROLLARY. *Each conjugate matrix R_i satisfies the minimum equation.*

If X is a matrix commutative with R_0, the matric equation $g(X) = 0$ can be written in the form $(X - R_0)(X - R_1)(X - R_2) \cdots (X - R_{m-1}) = 0$. This follows from Theorem 2 and the fact that the matrices R_i possess the commutative property for multiplication. Obviously, if $g_i(x) = 0$ is the

minimum equation of M_i, then $g_i(x)$ divides $g(x)$. Moreover, each of the first $\pi_q - 1$ conjugates will have $g(x) = 0$ for its minimum equation.

THEOREM 3. *If $g(x) = 0$ is of degree m and has the distinct roots $\rho_1, \rho_2, \cdots, \rho_q$ of the same multiplicity π, then the set of matrices $R_0, R_1, R_2, \cdots, R_{m-1}$ has the following properties:*

(a) *If* $\quad R_1 = \Theta(R_0) \equiv \xi_0 I + \xi_1 R_0 + \cdots + \xi_{m-1} R_0^{m-1}$

then $\quad R_i = \Theta(R_{i-1}) = \Theta^i(R_0), \quad (i = 1, 2, \cdots, \pi - 1),$

and $\quad R_0 = \Theta^\pi(R_0), \quad$ *where* $\Theta^2(R_0) \equiv \Theta(\Theta(R_0))$, *etc.**

(b) *If* $R_{l\pi} = \Theta_l(R_0) \equiv \xi_0^{(l)} I + \xi_1^{(l)} R_0 + \cdots + \xi_{m-1}^{(l)} R_0^{m-1}$

then $\quad R_{l\pi+1} = R_{l\pi+2} = \cdots = R_{(l+1)\pi-1} = \Theta_l(R_0)$

and $\quad R_{l\pi} = \Theta_1(R_{(l-1)\pi}) = \Theta_1^{\,l}(R_0), \quad (l = 1, 2, \cdots, q - 1).$

For the proof of (a) we note that, since each root has multiplicity π, $m = \pi q$. Let ω represent a primitive π-th root of unity. Since

$$R_1 = \Theta(R_0) \equiv \xi_0 I + \xi_1 R_0 + \cdots + \xi_{m-1} R_0^{m-1},$$

the conditions (11) become

$$(13) \qquad \begin{cases} \sum_{j=0}^{m-1} \xi_j \rho_k^{\,j} = \rho_k \\ \sum_{1}^{m-1} \xi_j \binom{j}{1} \rho_k^{\,j-1} = \omega \\ \sum_{n}^{m-1} \xi_j \binom{j}{n} \rho_k^{\,j-n} = 0, \qquad (n = 2, 3, \cdots, \pi - 1). \end{cases}$$

where $k = 1, 2, \cdots, q$. By the use of (6) we see that $R_i^{\,n}$ will have its k-th block of the form $(\rho_k^{\,n} I + \sum_{j=1}^{n} \binom{n}{j} \rho_k^{\,n-j} \omega^{ij} J^j)_s$, when the k-th block of R_i is of the form $(\rho_k I + \omega^i J)_s$, or of the form $(\rho_a^{\,n} I)$, when its k-th block has the form $(\rho_a I)$. It follows that the k-th block of $\Theta(R_i)$ has the form

$$(14) \qquad \Big(\sum_{j=0}^{m-1} \xi_j \rho_k^{\,j} I + \sum_{n-1}^{\pi-1} \{ \sum_{j=n}^{m-1} \xi_j \binom{j}{n} \omega^{in} \rho_k^{\,j-n} \} J^n \Big)_s \qquad \text{or}$$

$$(14') \qquad \Big(\sum_{j=0}^{m-1} \xi_j \rho_k^{\,j} I \Big).$$

---

* Compare with Pierce, *Bulletin of American Mathematical Society*, Vol. 36 (1930), p. 262.

The expression (14) can be written as

$$(15)\qquad \Big(\sum_{j=0}^{m-1} \xi_j \rho_k{}^j I + \sum_{n=1}^{\pi-1} \Big\{ \sum_{j=n}^{m-1} \xi_j \binom{j}{n} \rho_k{}^{j-n} \Big\} \omega^{in} J^n\Big)_s.$$

From (13) we see that (15) reduces to $(\rho_k I + \omega^{i+1} I)_s$ and (14′) reduces to $(\rho_k I)$. Thus the k-th block of $\Theta(R_i)$ is identical with the k-th block of R_{i+1}, for $i = 1, 2, \cdots, \pi - 1$. It follows that $R_{i+1} = \Theta(R_i)$. Therefore, $R_1 = \Theta(R_0)$, $R_2 = \Theta(R_1) = \Theta(\Theta(R_0)) \equiv \Theta^2(R_0)$, etc. Finally, since $\omega^\pi = 1$, $R_0 = \Theta^\pi(R_0)$.

The first statement of (b), that $R_{l\pi+i} = R_{l\pi}$ for $i = 1, 2, \cdots, p - 1$, follows immediately from the method of forming the conjugates. Since

$$R_\pi = \Theta_1(R_0) \equiv \xi_0{}^{(1)} I + \xi_1{}^{(1)} R_0 + \cdots + \xi_{m-1}^{(1)} R_0{}^{m-1},$$

the equations (11) become

$$(16)\qquad \begin{cases} \displaystyle\sum_{j=0}^{m-1} \xi_j{}^{(1)} \rho_k{}^j = \rho_{k+1} \\ \displaystyle\sum_{j=n}^{m-1} \xi_j{}^{(1)} \binom{j}{n} \rho_k{}^{j-1} = 0, \quad (n = 1, 2, \cdots, \pi - 1;\ k = 1, 2, \cdots, q). \end{cases}$$

If we form $\Theta_1(R_{l\pi})$, and use (16), it is easily seen that $R_{(l+1)\pi} = \Theta_1(R_{l\pi})$ and thus $R_{l\pi} = \Theta_1{}^l(R_0)$, $(l = 1, 2, \cdots, q - 1)$.

COROLLARY. *If $\pi = 1$ and $R_1 = \Theta(R_0)$, then $R_i = \Theta^i(R_0)$, $(i = 1, 2, \cdots, m - 1)$, and $R_0 = \Theta^m(R_0)$.*

THEOREM 4. *If $(x - r)^p = 0$ is the minimum equation of R_0, the conjugates to R_0 with respect to the minimum equation form, with R_0, the unique set of linear polynomials in R_0 such that*

(a) *They satisfy the minimum equation.*

(b) *Their elementary symmetric functions are equal to the elementary symmetric functions of the roots of $(x - r)^p = 0$, when these latter functions are considered as scalar matrices.*

This form of the minimum equation arises when R_0 has at least one block $(rI + J)_p$ and all other blocks are either $(rI + J)_k$, $k < p$ or (rI). Let T_i, $(i = 1, 2, \cdots, p - 1)$ represent matrices which satisfy the conditions of the theorem. Then

(17) $T_i = \phi_i(R_0) \equiv \xi_i I + \eta_i R_0$ and

(18) $(\xi_i I + \eta_i R_0 - rI)^p \equiv (\eta_i R_0 - \{r - \xi_i\} I)^p = 0$. Also,

(19) $S_k = \binom{p}{k} r^k I$, where S_k represents the k-th elementary symmetric function of the T_i.

If $\eta_i \neq 0$, $\{R_0 - [(r - \xi_i)/\eta_i] I\}^p = 0$. But, since $(x - r)^p = 0$ is the minimum equation of R_0, $(R_0 - rI)^p = 0$. Therefore,

$$\xi_i = r(1 - \eta_i). \tag{20}$$

If $\eta_i = 0$, obviously $\xi_i = r$. Therefore (20) holds for all cases.

Let σ_k be the k-th elementary symmetric function of the η_i, and define $\xi_0 = 0$, $\eta_0 = 1$, and $T_0 = \xi_0 I + \eta_0 R_0$. Then the T_i, $(i = 0, 1, 2, \cdots, p-1)$, are matrices of the type discussed in § 2. Consider any block of $T_0 \equiv R_0$ which has the form $(\{\xi_0 + \eta_0 r\} I + \eta_0 J)_p$ and its corresponding blocks $(\{\xi_i + \eta_i r\} I + \eta_i J)_p$ of the T_i, $(i = 1, 2, \cdots, p-1)$. The k-th elementary symmetric function of these p blocks will have the form $(b_{k0} I + \sum_{j=1}^{k} b_{kj} J^j)_p$, where $b_{k0} = \sum \prod_{i=0}^{k-1} (\xi_i + \eta_i r)$, $b_{kk} = \sigma_k$, and the remaining b_{kj} are symmetric functions of the ξ_i and η_i. Since $S_k = \binom{p}{k} r^k I$, $\sigma_k = 0$, $(k = 1, 2, \cdots, p-1)$. By use of the relation (20) and the fundamental theorem of symmetric functions, all terms of b_{kj} $(j = 1, 2, \cdots, k-1)$, and all terms except the first one of b_{k0} can be expressed as polynomials without a constant term in the σ_k and are therefore zero. From the relations $\sigma_k = 0$, it is evident that the η_i must be the roots of the equation $x^p - a = 0$. But $\eta_1 = 1$ and therefore $a = 1$. It follows that the η_i are the p-th roots of unity. Let ω be a primitive p-th root of unity and let $\eta_i = \omega^i$ $(i = 1, 2, \cdots, p-1)$. Then $\xi_i = r(1 - \omega^i)$ and $T_i = r(1 - \omega^i) I + \omega^i R_0$. It follows that $T_i \equiv R_i$.

5. *Conjugate matrices as the basis for a linear algebra.* The investigation of the conditions under which R_0 and its conjugates, as defined in § 3, form a basis for a linear algebra leads to the conclusion that they can form a set of basal elements for very special forms of the minimum equation.

THEOREM 5. *If the minimum equation, $g(x) = 0$, of R_0 has distinct roots, the matrices R_i form the basis for a linear algebra in the field of the roots r_k $(k = 1, 2, \cdots, m)$, under the conditions*

$$\sum_{k=1}^{m} r_k \omega_j^{m-k} \neq 0, \qquad (j = 0, 1, \cdots, m-1), \tag{21}$$

where the ω_j are the m-th roots of unity with $\omega_0 \equiv 1$.

Under the conditions on the roots

$$R_i = [(r_{i+1} I)_{p_1}, (r_{i+2} I)_{p_2}, \cdots, (r_{i+m} I)_{p_m}], \qquad (i = 0, 1, \cdots, m-1).$$

If the R_i form a set of basal elements, any element X of the algebra must have the form

$$X = \sum_{k=0}^{m-1} \xi_k R_k. \tag{22}$$

Obviously, the sum of any two elements and the product of any element by a scalar will be in the set of elements of the form (22). In order to investigate the product of any two elements it is sufficient to determine the conditions imposed by requiring that the product R_iR_j be of the form (22). These conditions are

$$\sum_{k=0}^{m-1} \xi_k r_{k+h} = r_{i+h} r_{j+h}, \qquad (h = 1, 2, \cdots, m), \tag{23}$$

in which the subscripts are considered as reduced modulo m. The determinant of the equations (23) is the circulant whose value is

$$\Delta' = (-1)^m \prod_{j=0}^{m-1} \left(\sum_{k=1}^{m} \omega_j{}^{m-k} r_k \right). \tag{24}$$

Under the conditions (21), $\Delta' \neq 0$ and the coefficients ξ_k can be determined.

THEOREM 6. *Let $g(x) = 0$, an equation of degree m, be the minimum equation of R_0 and let its distinct roots be $r_1, r_2, \cdots, r_{m-1}$, where r_1 has multiplicity 2 and each of the other r_i has multiplicity 1. Then the R_i form a set of basal elements for a linear algebra if*

$$\Delta'' \equiv \begin{vmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ r_1 & r_1 & r_2 & \cdots & r_{m-2} & r_{m-1} \\ r_2 & r_3 & r_4 & \cdots & r_1 & r_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{m-1} & r_1 & r_1 & \cdots & r_{m-3} & r_{m-2} \end{vmatrix} \neq 0. \tag{25}$$

Under the conditions on the roots,

$$R_0 = \{[(r_1I + J)_2]_{p_0}, (r_1I)_{p_1}, (r_2I)_{p_2}, \cdots, (r_{m-1}I)_{p_{m-1}}\}$$

where $p_0 \neq 0$, but the other p_i may be zero. Also

$$R_1 = \{[(r_1I - J)_2]_{p_0}, (r_1I)_{p_1}, (r_3I)_{p_2}, \cdots, (r_1I)_{p_{m-1}}\}$$

and

$$R_i = [(r_iI)_{2p_0}, (r_iI)_{p_1}, (r_{i+2}I)_{p_2}, \cdots, (r_{i-1}I)_{p_{m-1}}], \quad (i = 2, 3, \cdots, m-1).$$

As in Theorem 5, the product is the operation which imposes conditions on the roots r_k. If $R_iR_j = \sum_{k=0}^{m-1} \eta_k R_k$, the condition that the determinant Δ'' of

the coefficients of the η_k be different from zero is the condition (25) stated in the theorem.

THEOREM 7. *If the minimum equation, $g(x) = 0$, fails to satisfy the conditions of Theorems 5 and 6 as to the multiplicities of the roots, the set of matrices R_i cannot be a set of basal elements for a linear algebra.*

Case 1. Let $g(x) = 0$ be of degree m and have n distinct roots of multiplicity 2 and $m - 2n$ distinct roots of multiplicity 1, where $1 < n \leq m/2$.

Under these restrictions R_0 possesses at least two blocks, $(r_hI + J)_2$ and $(r_iI + J)_2$, $r_h \neq r_i$, and may possess blocks (r_jI). Accordingly, R_1 will have the corresponding blocks $(r_hI - J)_2$, $(r_iI - J)_2$ and (r_jI) or $(r_{j+1}I)$ (according as $j \leq n$ or $j > n$). The remaining R_k, $(k = 2, 3, \cdots, m-1)$, will consist of blocks (r_lI). Moreover, $m \geq 4$. The attempt to express R_0R_k in the form $\sum \xi_iR_i$ leads to conditions which include

$$(26) \qquad \xi_0 - \xi_1 = r_k, \quad \xi_0 - \xi_1 = r_{k+1}, \cdots, \quad \xi_0 - \xi_1 = r_{k+n-1}.$$

Since $n \geq 2$, there are at least two of these equations. The r_k are distinct so that the conditions (26) are incompatible. Similarly $R_1R_k = \sum \xi_iR_i$ leads to a set of incompatible conditions.

Case 2. Let $g(x) = 0$ have a root r_1 of multiplicity $p > 2$. Then R_0 has at least one block of the form $(r_1I + J)_p$. The R_i, $(i = 1, 2, \cdots, p-1)$, will have the corresponding block $(r_1I + \omega^iJ)_p$ and the remaining R_k, $(k = p, p+1, \cdots, m-1)$, will have $(r_lI)_p$, $r_l \neq r_1$. The product R_iR_j, $(i, j = 0, 1, 2, \cdots, p-1)$, has the corresponding block

$$(r_1^2I + \{\omega^i + \omega^j\}\ rJ + \omega^{i+j}J^2)_p.$$

If we attempt to express R_iR_j as $\sum \xi_kR_k$ we are led to the condition $\omega^{i+j} = 0$. Since ω is a p-th root of unity, this condition cannot be satisfied.

THEOREM 8. *If R_0 is a matrix which satisfies either set of conditions that the R_i form a set of basal elements for a linear algebra, the algebra will contain a principal unit.*

Case 1. Let $g(x) = 0$ have distinct roots. If there exists an element $\sum_{i=0}^{m-1} \alpha_iR_i$ such that $\sum_{i=0}^{m-1} \alpha_iR_iR_j = R_j$, the following conditions on the α_i must be satisfied:

$$(27) \qquad \sum_{i=0}^{m-1} \alpha_i r_{i+k} r_{j+k} = r_{j+k}, \qquad (k = 1, 2, \cdots, m).$$

If the r_i are all different from zero, (27) reduces to

$$(28) \qquad \sum_{i=0}^{m-1} \alpha_i r_{i+k} = 1.$$

The determinant of the coefficients of these equations (28) is Δ' given by (24). Since the conditions of Theorem 5 are satisfied, $\Delta' \neq 0$, and the equations (28) can be solved for the α_i. If any root, say r_m, is zero then the m equations (27) will reduce to the first $m - 1$ of the equations (28). Since the rank of the matrix of the coefficients of these equations is $m - 1$, there will be a single infinity of solutions but only one independent solution.

Case 2. Let one root of $g(x) = 0$ have multiplicity 2 and the others have multiplicity 1. By an argument similar to that used above, it can be shown that the determinant of the coefficients is Δ'', given by (25).

6. *Discussion of other sets of conjugate matrices.* There exist other sets of matrices conjugate to R_0 with respect to its minimum equation, $g(x) = 0$, which possess the properties (i) and (ii). If a set of matrices S_i, $(i = 1, 2, \cdots, m - 1)$, is to possess the property (i), it is evident that the S_i must have the form

$$[(d_{i0}^{(1)}I + \sum_{k=1}^{p_1-1} d_{ik}^{(1)}J^k)_{p_1}, (d_{i0}^{(2)}I + \sum_{k=1}^{p_2-1} d_{ik}^{(2)}J^k)_{p_2}, \cdots,$$
$$(d_{i0}^{(a)}I + \sum_{k=1}^{p_a-1} d_{ik}^{(a)}J^k)_{p_a}, (d_{i0}^{l}I), \cdots, (d_{i0}^{(c)}I)].$$

The conjugates are, therefore, matrices of the type discussed in § 2, and the property (ii) can be discussed by investigating a typical block and the corresponding block in each conjugate. Let $(d_{00}I + \sum_{k=1}^{p-1} d_{0k}J^k)_p$ represent any block of R_0, where $d_{00} = r$, $d_{01} = 1$, and the $d_{0k} = 0$ for $k = 2, 3, \cdots, p - 1$. The corresponding block of S_i, $(i = 1, 2, \cdots, m - 1)$, is $(d_{i0}I + \sum_{k=1}^{p-1} d_{ik}J^k)_p$. From (ii) it follows that the elementary symmetric functions of the d_{i0} must be equal to the elementary symmetric functions of the roots of $g(x) = 0$. It follows that the d_{i0} must take on the values of the roots of $g(x) = 0$ in some order. Moreover, the coefficients of J^k appearing in these elementary symmetric functions, must all be zero. The d_{i1} cannot all be zero unless every root has multiplicity 1, but the d_{ik} for $k \neq 0$ or 1 may be zero.

Even when the $d_{i1}^{(k)}$ are chosen as the π_k-th roots of unity or zero and the $d_{ij}^{(k)}$ $(j > 1)$ are chosen as zero, there remains a choice in the arrangement of the conjugates of the blocks in the formation of the conjugate matrices. If a knowledge of the actual form of the minimum equation of

each conjugate is desirable, the conjugates of the blocks can be so combined that the minimum equation is immediately evident. One such set of conjugate matrices is described in the following paragraph.

The first $\pi_1 - 1$ conjugates are formed by replacing each block of the form $(\rho_1 I + J)_{p_i}$ successively by the blocks $(\rho_1 I + \omega_{\pi_1}{}^j J)_{p_i}$, $(j = 1, 2, \cdots, \pi_1 - 1)$, and every other block by $(\rho_1 I)_{p_k}$. The next $\pi_2 - 1$ conjugates are formed by replacing each block $(\rho_2 I + J)_{p_i}$ by the blocks $(\rho_2 I + \omega_{\pi_2}{}^j J)_{p_i}$, $(j = 1, 2, \cdots, \pi_2 - 1)$, etc. This process defines $m - q$ conjugates. The remaining conjugates are formed by replacing each block containing ρ_i by $(\rho_{i+k} I)_{p_j}$, $(k = 1, 2, \cdots, q - 1)$, where the subscript on the ρ is to be reduced modulo q. Obviously, $(x - \rho_1)^{\pi_1} = 0$ will be the minimum equation for the first $\pi_1 - 1$ conjugates, $(x - \rho_2)^{\pi_2} = 0$ will be the equation for the next $\pi_2 - 1$, etc., and $\prod_{i=1}^{q} (x - \rho_i) = 0$ will be the minimum equation for the last $q - 1$ conjugates. Every conjugate will satisfy $g(x) = 0$, but no conjugate will have $g(x) = 0$ as its minimum equation unless $g(x) \equiv (x - \rho)^{\pi}$ or $g(x) \equiv \prod_{i=1}^{q} (x - \rho_i)$. In either of these cases the conjugates are the same by this method of formation as by the method of § 3.

RELATIONS BETWEEN THE PROJECTIVE AND METRIC DIFFERENTIAL GEOMETRIES OF SURFACES.*

By O. W. Albert.

1. *Introduction.* In the First Memoir of "Projective Differential Geometry of Curved Surfaces," by Professor E. J. Wilczynski, *Transactions of the American Mathematical Society,* Vol. 8, the projective geometry of a surface is based on a completely integrable system of two linear homogeneous partial differential equations of second order in one dependent and two independent variables. When the surface is non-degenerate, non-developable, and is referred to its asymptotic lines as parametric curves, these equations reduce to the intermediate form:

$$(1)\qquad \begin{aligned} y_{uu} + 2ay_u + 2by_v + cy &= 0, \\ y_{vv} + 2a'y_u + 2b'y_v + c'y &= 0. \end{aligned}$$

Such a system has four linearly independent solutions: $y^{(1)}, y^{(2)}, y^{(3)}, y^{(4)}$. When these solutions are interpreted as the four homogeneous coördinates of a point P_y in space, its locus is an integral surface S of equations (1). The most general system of solutions of these equations has the form:

$$(2)\qquad \eta_i = \sum_{k=1}^{4} c_{ik} y_k,$$

where

$$| c_{ik} | \neq 0; \qquad (i = 1, 2, 3, 4).$$

Therefore the most general integral surface of equations (1) is a projective transformation of any particular one. Any projective property of the integral surface, independent of the special analytic method of representation, will then be given by an invariant equation or system of equations. Such equations, involving the coefficients of (1) and their derivatives, remain invariant for all transformations of the form:

$$(3)\qquad y = \lambda(u, v)\bar{y}, \quad \bar{u} = \alpha(u), \quad \bar{v} = \beta(v).$$

Other configurations related to the integral surface will be given by the covariants.

In his paper on "Relations between Projective and Metric Differential Geometry," *American Journal of Mathematics,* Vol. 39, F. M. Morrison has

* Presented to the Society, November 28, 1931.

identified Cartesian coördinates of a point on the surface with the solutions of (1) by making $y^{(4)} = 1$, which reduces (1) to:

$$
(4) \qquad \begin{aligned} y_{uu} + 2ay_u + 2by_v &= 0, \\ y_{vv} + 2a'y_u + 2b'y_v &= 0. \end{aligned}
$$

By comparing these equations with the Gauss equations of the metric theory of surfaces,

$$
(5) \qquad \begin{aligned} x_{uu} &= \begin{Bmatrix} 11 \\ 1 \end{Bmatrix} x_u + \begin{Bmatrix} 11 \\ 2 \end{Bmatrix} x_v + DX, \\ x_{uv} &= \begin{Bmatrix} 12 \\ 1 \end{Bmatrix} x_u + \begin{Bmatrix} 12 \\ 2 \end{Bmatrix} x_v + D'X, \\ x_{vv} &= \begin{Bmatrix} 22 \\ 1 \end{Bmatrix} x_u + \begin{Bmatrix} 22 \\ 2 \end{Bmatrix} x_v + D''X, \end{aligned}
$$

Morrison observed that if $D = D'' = 0$ so that $u = c$ and $v = c$ become the asymptotic lines, then the first and third equations of (5) have exactly the form of equations (4). This enabled him to express the coefficients of (4) in terms of the Christoffel symbols in (5) and to determine transformations from the homogeneous coördinate system with a semi-covariant tetrahedron of reference to a rectangular Cartesian system. In this way a metric study was made of the two osculating linear complexes of the asymptotic curves, which had been investigated projectively by Professor Wilczynski.

It is our purpose in this paper to apply a new method for studying relations between the projective and metric differential geometries of surfaces. We will use Morrison's equations (11) of the paper mentioned above, which express the transformations from a homogeneous coördinate system with a semi-covariant tetrahedron of reference to a rectangular Cartesian system. However, instead of using a moving trihedral which makes a rectangular coördinate system at a point of a surface we shall use one which makes an oblique coördinate system, namely the surface normal and the tangents to the asymptotic curves. By transforming from rectangular axes to these special oblique axes, we can obtain equations expressing the relations between the homogeneous coördinates and the Cartesian coördinates of the same point referred to our special trihedral. We will show that these equations give in simpler form all the results Morrison obtained in his metric study of the two osculating linear complexes of the asymptotic curves. We will then investigate metrically the two osculating linear complexes of the osculating ruled surfaces, the six congruences determined by pairs of these four linear complexes, and relations between the four complexes. All these complexes

have been discussed projectively by Professor E. J. Wilczynski in his Second Memoir (*Transactions of the American Mathematical Society,* Vol. 9), but up to this time the complicated symbolism of Morrison has discouraged any attempt to add to the results he obtained metrically.

2. *Fundamental transformations of coördinates.* Before obtaining the fundamental equations desired we will outline briefly Morrison's derivation of equations (11) of his paper.

Assuming that the integrability conditions of system (4) are satisfied so the system has four linear independent solutions $y^{(1)}$, $y^{(2)}$, $y^{(3)}$, $y^{(4)} = 1$, and that $D = D'' = 0$ so the lines $u = c$ and $v = c$ become the asymptotic curves, then by comparing (4) with the first and third of equations (5) Morrison found:

$$(6)\qquad \begin{aligned} a &= -(1/2)\left\{\begin{matrix}11\\1\end{matrix}\right\}; & b &= -(1/2)\left\{\begin{matrix}11\\2\end{matrix}\right\}; \\ a' &= -(1/2)\left\{\begin{matrix}22\\1\end{matrix}\right\}; & b' &= -(1/2)\left\{\begin{matrix}22\\2\end{matrix}\right\}. \end{aligned}$$

The four fundamental semi-covariants of system (1) are:

$$(7)\qquad \begin{aligned} &y,\ z = y_u + ay, \qquad \rho = y_v + b'y, \\ &\sigma = y_{uv} + b'y_u + ay_v + (1/2)(a_v + b'_u + 2ab')y. \end{aligned}$$

By substituting the four linear independent solutions $y^{(1)}$, $y^{(2)}$, $y^{(3)}$, 1 for y in (7); by using the resulting four semi-covariant points, as did Professor Wilczynski, for vertices of a tetrahedron of reference; by letting $y^{(1)} = x$, $y^{(2)} = y$, $y^{(3)} = z$; and by determining x_{uv}, y_{uv}, z_{uv} from equations of the form of the second equation of (5), Morrison found the following relations between the rectangular Cartesian coördinates $\bar{\xi}, \bar{\eta}, \bar{\zeta}$ and the homogeneous coördinates x_1, x_2, x_3, x_4 of the same point:

$$(8)\qquad \begin{aligned} \bar{\omega}\bar{\xi} &= x_1 x + x_2(x_u + ax) + x_3(x_v + b'x) \\ &\qquad + x_4[(b' - 2c)x_u + (a - 2d)x_v + Rx + D'X], \\ \bar{\omega}\bar{\eta} &= x_1 y + x_2(y_u + ay) + x_3(y_v + b'y) \\ &\qquad + x_4[(b' - 2c)y_u + (a - 2d)y_v + Ry + D'Y], \\ \bar{\omega}\bar{\zeta} &= x_1 z + x_2(z_u + az) + x_3(z_v + b'z) \\ &\qquad + x_4[(b' - 2c)z_u + (a - 2d)z_v + Rz + D'Z], \\ \bar{\omega} &= x_1 + x_2 a + x_3 b' + x_4 R, \end{aligned}$$

where

$$(9)\qquad \begin{aligned} &c = -(1/2)\left\{\begin{matrix}12\\1\end{matrix}\right\}, \qquad d = -(1/2)\left\{\begin{matrix}12\\2\end{matrix}\right\}; \\ &R = (1/2)(a_v + b'_u + 2ab'), \end{aligned}$$

and where X, Y, Z are the direction cosines of the surface normal. Since the determinant of the right members of (8) $= -HD'$ and is $\neq 0$ for any non-developable surface, it was possible for Morrison to solve equations (8) for the homogeneous coördinates x_1, x_2, x_3, x_4, giving his equations (11), which are the inverse of transformations (8).

In order to transform from rectangular axes to our special oblique axes and also to translate axes to the general surface point x, y, z, we get by standard formulae:

$$\begin{aligned}\bar{\xi} &= (x_u/E^{1/2})\xi + (x_v/G^{1/2})\eta + X\zeta + x\\ \bar{\eta} &= (y_u/E^{1/2})\xi + (y_v/G^{1/2})\eta + Y\zeta + y\\ \bar{\zeta} &= (z_u/E^{1/2})\xi + (z_v/G^{1/2})\eta + Z\zeta + z,\end{aligned} \tag{10}$$

where the coefficients of ξ are direction cosines of the tangent to $v = c$, those of η are direction cosines of the tangent to $u = c$, and those of ζ are the same as in (8). Substituting (10) in Morrison's equations (11) and letting $\bar{\omega}/H^2 = \rho$ be the new proportionality factor, we find:

$$\begin{aligned}\rho x_1 &= (D'a/E^{1/2})\xi + (D'b'/G^{1/2})\eta + S\zeta - D',\\ \rho x_2 &= (-D'/E^{1/2})\xi + (b' - 2c)\zeta,\\ \rho x_3 &= (-D'/G^{1/2})\eta + (a - 2d)\zeta,\\ \rho x_4 &= -\zeta,\end{aligned} \tag{11}$$

where

$$S = (1/2)(a_v + b'_u - 2ab') + 2(ac + b'd). \tag{12}$$

These are the fundamental equations desired, for they express in simple form the relations between the homogeneous coördinates of a point referred to the semi-covariant tetrahedron and the Cartesian coördinates of the same point referred to our special trihedral.

3. *The osculating linear complexes of the asymptotic curves.* In this section we will rewrite enough of Morrison's results in our special coördinate system to show that our equations (11) will give in simpler form all the results he obtained.

The projective equations of these complexes of $v = c$ and $u = c$ respectively, were found by Professor Wilczynski to be:

$$C', \quad -b_v\bar{\omega}_{34} - b\bar{\omega}_{14} + b\bar{\omega}_{23} = 0; \tag{13}$$

$$C'', \quad -a'_u\bar{\omega}_{42} + a'\bar{\omega}_{14} + a'\bar{\omega}_{23} = 0, \tag{14}$$

where $\bar{\omega}_{ik}$ are the Plücker homogeneous line coördinates.

From equations (11) the transformations of the line coördinates $\bar{\omega}_{ik}$ in terms of the ω_{ik} for our special Cartesian system become:

$$
(15)\quad
\begin{aligned}
&\bar\omega_{34}=(D'/G^{1/2})\omega_{23},\qquad \bar\omega_{42}=(D'/E^{1/2})\omega_{31},\\
&\bar\omega_{14}=(D'a/E^{1/2})\omega_{31}-(D'b'/G^{1/2})\omega_{23}-D'\omega_{34},\\
&\bar\omega_{23}=[D'^2/(EG)^{1/2}]\omega_{12}+[(b'-2c)D'/G^{1/2}]\omega_{23}+[(a-2d)D'/E^{1/2}]\omega_{31}.
\end{aligned}
$$

By (15) the equation of C' becomes:

$$(16)\qquad bD'\omega_{12}-2bdG^{1/2}\omega_{31}+\alpha E^{1/2}\omega_{23}+b(EG)^{1/2}\omega_{34}=0,$$

where

$$\alpha=2bb'-2bc-b_v.$$

In a similar way the equation (14) of C'' becomes:

$$(17)\qquad a'D'\omega_{12}+\beta G^{1/2}\omega_{31}-2a'cE^{1/2}\omega_{23}-a'(EG)^{1/2}\omega_{34}=0,$$

where

$$\beta=2aa'-2a'd-a'_u.$$

At once we notice that the notation of equations (16) and (17) is considerably simpler than that of equations (19) and (29) of Morrison in the paper mentioned above. This property has more significance in the computations of the analysis which Morrison made from these equations. In fact, the writer has completed this analysis in detail by using equations (16) and (17), but will give here only a summary of the reductions of these two complexes to their simplest forms, the derivations of simpler equations for the directrices of the congruence they determine, and two corrections to Morrison's paper, as it appears in the reference cited above.

The equation of any linear complex, where ω_{ik} are defined for a Cartesian system of axes, can be written thus:

$$(18)\qquad a_{12}\omega_{12}+a_{31}\omega_{31}+a_{23}\omega_{23}+a_{14}\omega_{14}+a_{24}\omega_{24}+a_{34}\omega_{34}=0.$$

The axis of the complex of this form would have the following equation for a rectangular coördinate system, as given by Plücker's discussion in his "Neue Geometrie des Raumes," page 32:

$$
(19)\quad
\begin{aligned}
&[\xi-(a_{24}a_{12}-a_{34}a_{31})/(a_{23}^2+a_{31}^2+a_{12}^2)]/a_{23}\\
&\quad=[\eta-(a_{34}a_{23}-a_{14}a_{12})/(a_{23}^2+a_{31}^2+a_{12}^2)]/a_{31}\\
&\quad=[\zeta-(a_{14}a_{31}-a_{24}a_{23})/(a_{23}^2+a_{31}^2+a_{12}^2)]/a_{12}.
\end{aligned}
$$

When this equation is derived for our special oblique system of axes, it becomes:

$$
(20)\quad
\begin{aligned}
&[\xi-(a_{24}a_{12}-a_{34}a_{31}-a_{34}a_{23}\cos\omega)/\Sigma]/a_{23}\\
&\quad=[\eta-(a_{34}a_{23}-a_{14}a_{12}+a_{34}a_{31}\cos\omega)/\Sigma]/a_{31}\\
&\quad=\{\zeta-[a_{14}a_{31}-a_{24}a_{23}+(a_{14}a_{23}-a_{24}a_{31})\cos\omega]/\Sigma\}/a_{12},
\end{aligned}
$$

where

$$\Sigma=a_{23}^2+a_{31}^2+a_{12}^2+2a_{23}a_{31}\cos\omega,$$

where a_{23}, a_{31}, a_{12} are proportional to direction ratios, and ω is the angle between the non-rectangular axes. By substituting $\cos\omega = F/(EG)^{1/2}$ for asymptotic tangents in (20) the axis of C' becomes:

$$(21)\qquad [\xi - (2b^2dGE^{1/2} - b\alpha FE^{1/2})/T_1{}^2]/\alpha E^{1/2} = [\eta - (b\alpha EG^{1/2} - 2b^2dFG^{1/2})/T_1{}^2]/-2bdG^{1/2} = \zeta/bD',$$

where

$$T_1{}^2 = \alpha^2E + 4b^2d^2G + b^2D'^2 - 4bd\alpha F.$$

If the fixed point of the axis (19) for a rectangular system of axes, be substituted in (18) we get the equation of the polar plane associated with that point as pole to be:

$$(22)\qquad a_{23}\xi + a_{31}\eta + a_{12}\zeta = 0.$$

This is then that plane, of the system of parallel planes perpendicular to the axis, which passes through the surface point and which Morrison called the principal plane. It is easily seen that, if we put ω equal to ninety degrees, equation (20) reduces to (19) and hence we get (22).

To find the parameter of C' we will transform the axes to a rectangular system, whose origin is the intersection of the axis of the complex with the tangent plane, the ξ' axis being the line of intersection of the plane (22) with the tangent plane, and the ζ' axis being the axis of the complex. From equations (19), (22), and the standard formulae for transformation of axes, we find the corresponding transformations of the line coördinates. By substituting these in (16) we get the simplest form of the equation of the complex C', the coefficient of ω'_{34} being its parameter:

$$(23)\qquad \omega'_{12} + [b^2D'(EG)^{1/2}/T^2]\omega'_{34} = 0,$$

where

$$T^2 = \alpha^2E + 4b^2d^2G + b^2D'^2.$$

By a similar analysis of the complex C'' we find the simplest form of its equation to be:

$$(24)\qquad \omega_{12}'' - [a'^2D'(EG)^{1/2}/V^2]\omega_{34}'' = 0,$$

where

$$V^2 = \beta^2G + 4a'^2c^2E + a'^2D'^2.$$

From (23) and (24) we see that if the surface and the asymptotic curves are real, the osculating linear complexes of the asymptotic curves through the surface point are oppositely twisted, which agrees with Morrison's result.

If the complex C_λ of the pencil determined by C' and C'' is special, λ must satisfy the condition:

(25) $$a_{12}a_{34} + a_{31}a_{24} + a_{23}a_{14} = 0,$$

and hence $$\lambda = \pm b/a'.$$

If $\lambda = - b/a'$ the axis or directrix of the first kind becomes after substituting in (20) and simplifying:

(26) $$bG^{1/2}\beta_1\xi + a'E^{1/2}\alpha_1\eta = 2a'b(EG)^{1/2}, \quad \zeta = 0,$$

where $$\alpha_1 = \alpha + 2bc, \quad \beta_1 = \beta + 2a'd.$$

If $\lambda = b/a'$ the axis or directrix of the second kind becomes by (20):

(27) $$\xi/a'E^{1/2}\alpha_2 = \eta/bG^{1/2}\beta_2 = \zeta/2a'bD',$$

where $$\alpha_2 = \alpha - 2bc, \quad \beta_2 = \beta - 2a'd.$$

Equations (26) and (27) represent the directrices of the congruence determined by C' and C'' in forms which will give some interesting relations in another section of this paper.

The first of the two errors mentioned above is in the first equation of (40), Section IV of Morrison's paper. The constant term in the left member of this equation should be:

$$-H(b - \lambda a'), \text{ not } + H(b - \lambda a'), \text{ and not } 0, \text{ as in one edition.}$$

The second error occurs in part b of Section VII. The right member of the reduced equation of the cylindroid should contain the factor H in the numerator of the coefficient.

4. *The osculating linear complexes of the osculating ruled surfaces.* This and the remaining sections are new, except where definite references are given. Where any results appear incidentally which were already known by the projective theory, this fact has been noted to show the agreement in the results obtained projectively and metrically. All results stated as theorems are new.

The equations of these complexes, referred to the semi-covariant tetrahedron, were found and studied projectively by Professor Wilczynski. These equations for ruled surfaces of the first and second kinds, R_1 and R_2 respectively, were given in the forms:

(28) $$C_1, \; a_{12}\bar{\omega}_{12} + a_{13}\bar{\omega}_{13} + a_{23}\bar{\omega}_{23} + a_{14}\bar{\omega}_{14} + a_{42}\bar{\omega}_{42} + a_{34}\bar{\omega}_{34} = 0;$$

(29) $$C_2, \; b_{12}\bar{\omega}_{12} + b_{13}\bar{\omega}_{13} + b_{23}\bar{\omega}_{23} + b_{14}\bar{\omega}_{14} + b_{42}\bar{\omega}_{43} + b_{34}\bar{\omega}_{34} = 0;$$

where

$$a_{12}=0,\quad a_{13}=2^8a'^2C,\quad a_{23}=-2^3(a'\theta_v-2a'_v\theta-2^4a'a'_uC)=a_{14},$$
$$a_{42}=-[C(\theta+64a'^2_u)+16\theta a'_ua'_v/a'-8a'_u\theta_v],\quad a_{34}=-2^9a'^3bC;$$
$$C=2^3(a'_{uv}-a'_ua'_v/a'-4a'^2b),\quad \theta=\Theta_4 \text{ invariant of } R_1;$$
$$b_{12}=2^8b^2C',\quad b_{13}=0,\quad -b_{23}=-2^3(b\theta'_u-2b_u\theta'-2^4bb_vC')=b_{14},$$
$$b_{42}=2^9a'b^3C',\quad b_{34}=C'(\theta'+64b_v^2)+16\theta'b_ub_v/b-8b_v\theta'_u;$$
$$C'=2^3(b_{uv}-b_ub_v/b-4a'b^2),\quad \theta'=\Theta'_4 \text{ invariant of } R_2.$$

By equations (15), the equation (28) of C_1 becomes:

$$\begin{aligned}(30)\quad & a'[2^7\bar{C}(\beta+2a'd)+a'\theta_v-2a'_v\theta]D'\omega_{12}\\ &+[2^6\bar{C}(\beta^2-4a'^2d^2)+\beta(a'\theta_v-2a'_v\theta)+\bar{C}\theta]G^{1/2}\omega_{31}\\ &-2a'[2^7\bar{C}c(\beta+2a'd)+2^6\bar{C}a'(a_v+b'_u-4a'b)\\ &+c(a'\theta_v-2a'_v\theta)]E^{1/2}\omega_{23}\\ &+2^8a'^2\bar{C}D'E^{1/2}\omega_{24}-a'[2^7\bar{C}(\beta-2a'd)+a'\theta_v-2a'_v\theta](EG)^{1/2}\omega_{34}=0,\end{aligned}$$

where $$\bar{C}=a'C/2^3.$$

In addition to our former notation $\alpha_1, \beta_1, \alpha_2, \beta_2$ let:

$$\begin{aligned}& a'\theta_v-2a'_v\theta=\bar{A},\quad a_v+b'_u-4a'b=\bar{\bar{R}},\quad 2^7\bar{C}\beta_1+\bar{A}=\beta'_1,\\ (31)\quad & 2^7\bar{C}\beta_2+\bar{A}=\beta'_2,\quad 2^6\bar{C}\beta_1\beta_2+\beta\bar{A}+\bar{C}\theta=\beta',\\ & 2^7\bar{C}c\beta_1+2^6\bar{C}a'\bar{\bar{R}}+c\bar{A}=c'.\end{aligned}$$

Then the equation of C_1, the osculating linear complex of R_1, is:

$$\begin{aligned}(32)\quad & a'\beta'_1D'\omega_{12}+\beta'G^{1/2}\omega_{31}\\ &-2a'c'E^{1/2}\omega_{23}+2^8a'^2\bar{C}D'E^{1/2}\omega_{24}-a'\beta'_2(EG)^{1/2}\omega_{34}=0.\end{aligned}$$

We notice that, while the notation here is necessarily more complicated, this equation and equation (17) are very similar in form. We may call attention to the fact that both of these complexes are related to the same asymptotic curve through the surface point, $u=c$, and both are referred to our special trihedral. They coincide, if $\bar{C}=0$.

We will now reduce the equation of this complex to its simplest form by transforming the axes to a rectangular system, whose origin is the intersection of the axis of the complex with the tangent plane, whose ξ' axis coincides with the intersection of that polar plane of this complex given by (22) and the tangent plane, and whose ζ' axis is the axis of the complex. By equations (19) and (22) we get the point and lines desired.

The new origin becomes by (19):

$$\begin{aligned}(33)\quad & \xi_1=2^8a'\bar{C}E^{1/2}/\beta'_1-a'AB'GE^{1/2}/\beta'_1V'^2,\\ & \eta_1=-2a'^2c'AEG^{1/2}/\beta'_1V'^2,\quad \zeta_1=0,\end{aligned}$$

where

$$A = (a_{12}a_{34} + a_{31}a_{24})/a'^2D'(EG)^{1/2} = 2^8\bar{C}^2\theta - \bar{A}^2,$$
$$V'^2 = a_{23}^2 + a_{31}^2 + a_{12}^2 = 4a'^2c'^2E + \beta'^2G + a'^2\beta'_1{}^2D'^2.$$

The polar plane through the surface point is by (22):

$$(34) \qquad -2a'c'E^{1/2}\xi + \beta'G^{1/2}\eta + a'\beta'_1D'\zeta = 0.$$

The line of intersection of this plane with the tangent plane becomes:

$$(35) \qquad -2a'c'E^{1/2}\xi + \beta'G^{1/2}\eta = 0, \quad \zeta = 0.$$

Then the equations of transformation become:

$$(36) \qquad \begin{aligned} \xi &= (\beta'G^{1/2}/W')\xi' - (2a'^2c'\beta'_1D'E^{1/2}/V'W')\eta' - (2a'c'E^{1/2}/V')\zeta' + \xi_1, \\ \eta &= (2a'c'E^{1/2}/W')\xi' + (a'\beta'\beta'_1D'G^{1/2}/V'W')\eta' + (\beta'G^{1/2}/V')\zeta' + \eta_1, \\ \zeta &= (-W'/V')\eta' + (a'\beta'_1D'/V')\zeta', \end{aligned}$$

where
$$W'^2 = 4a'^2c'^2E + \beta'^2G,$$

and where the positive directions of the axes are arbitrary. Substituting the values of the corresponding transformations of line coördinates in equation (32), we find the equation of C_1 reduces to:

$$(37) \qquad \omega'_{12} + [a'^2AD'(EG)^{1/2}/V'^2]\omega'_{34} = 0.$$

Thus the parameter of the complex C_1 is:

$$(38) \qquad P'_1 = a'^2AD'(EG)^{1/2}/V'^2 = -a'^2D'(EG)^{1/2}\mathfrak{A}/2^6V'^2,$$

where $\mathfrak{A}$ is the invariant of C_1 in the form found by Professor Wilczynski. It follows that when $A = 0$ in (37), the complex is special, as it should be.

We will give in briefer form a similar analysis for the complex C_2, which osculates the ruled surface R_2 of the second kind. By equation (15), C_2 becomes:

$$(39) \qquad \begin{aligned} b(2^7\bar{C}'\alpha_1 + \bar{B})D'\omega_{12} &- 2b(2^7\bar{C}'d\alpha_1 + 2^6\bar{C}'b\bar{R} + d\bar{B})G^{1/2}\omega_{31} \\ &+ (2^6\bar{C}'\alpha_1\alpha_2 + \alpha B + \bar{C}'\theta')E^{1/2}\omega_{23} \\ &- 2^8b^2\bar{C}'D'G^{1/2}\omega_{14} + b(2^7\bar{C}\,\alpha_2 + \bar{B})(EG)^{1/2}\omega_{34} = 0, \end{aligned}$$

where
$$\bar{C}' = bC'/2^3, \qquad \bar{B} = b\theta'_u - 2b_u\theta'.$$

To get the form of this equation corresponding to (32), let:

$$(40) \qquad \begin{aligned} &2^7\bar{C}'\alpha_1 + \bar{B} = \alpha'_1, \qquad 2^7\bar{C}'d\alpha_1 + 2^6\bar{C}'b\bar{R} + d\bar{B} = d', \\ &2^6\bar{C}'\alpha_1\alpha_2 + \alpha\bar{B} + \bar{C}'\theta' = \alpha', \qquad 2^7\bar{C}\alpha_2 + \bar{B} = \alpha'_2. \end{aligned}$$

Then the equation of C_2 becomes:

(41) $b\alpha'_1 D'\omega_{12} - 2bd'G^{1/2}\omega_{31} + \alpha' E^{1/2}\omega_{23} - 2^8 b^2 \bar{C}' D' G^{1/2}\omega_{14} + b\alpha'_2 (EG)^{1/2}\omega_{34} = 0.$

By equations (19) and (22) we can transform the axes to a rectangular system similar to that used for C_1 and reduce C_2 to its simplest form. The new origin is:

(42) $\xi_2 = 2b^2 d' BGE^{1/2}/\alpha'_1 T'^2, \quad \eta_2 = bB\alpha' EG^{1/2}/\alpha'_1 T'^2 + 2^8 b\bar{C}' G^{1/2}/\alpha'_1, \quad \zeta_2 = 0,$

where $B = (b_{12}b_{34} + b_{23}b_{14})/b^2 D'(EG)^{1/2} = \bar{B}^2 - 2^8 \bar{C}'^2 \theta',$
$T'^2 = b_{23}^2 + b_{31}^2 + b_{12}^2 = \alpha'^2 E + 4b^2 d'^2 G + b^2 \alpha'^2_1 D'^2.$

The polar plane through the surface point is:

(43) $$\alpha' E^{1/2}\xi - 2bd'G^{1/2}\eta + b\alpha'_1 D'\zeta = 0.$$

Its intersection with the tangent plane is:

(44) $$\alpha' E^{1/2}\xi - 2bd'G^{1/2}\eta = 0, \quad \zeta = 0.$$

After transforming axes and line coördinates, (41) reduces to the simplest form of C_2:

(45) $$\omega_{12}'' + [b^2 BD'(EG)^{1/2}/T'^2]\omega_{34}'' = 0,$$

where the axes have been oriented in the same way as for equation (37) of C_1. Thus the parameter of C_2 is:

(46) $$P'_2 = b^2 BD'(EG)^{1/2}/T'^2 = - b^2 D'(EG)^{1/2}\mathfrak{B}/2^6 T'^2.$$

5. *Congruence determined by complexes C_1 and C_2.* The pencil of linear complexes determined by C_1 and C_2 is:

(47) $$D'(b\alpha'_1 + \lambda a'\beta'_1)\omega_{12} - G^{1/2}(2bd' - \lambda\beta')\omega_{31} + E^{1/2}(\alpha' - 2\lambda a'c')\omega_{23}$$
$$- 2^8 b^2 \bar{C}' D' G^{1/2}\omega_{14} + 2^8 \lambda a'^2 \bar{C} D' E^{1/2}\omega_{24} + (EG)^{1/2}(b\alpha'_2 - \lambda a'\beta'_2)\omega_{34} = 0.$$

The special complexes of the pencil are determined by (25) which gives:

(48) $$D'(b\alpha'_1 + \lambda a'\beta'_1)(EG)^{1/2}(b\alpha'_2 - \lambda a'\beta'_2) - G^{1/2}(2bd' - \lambda\beta')2^8\lambda a'^2 \bar{C} D' E^{1/2}$$
$$- E^{1/2}(\alpha' - 2\lambda a'c')2^8 b^2 \bar{C}' D' G^{1/2} = 0.$$

By simplifying, (48) reduces to:

(49) $$a'^2 A\lambda^2 + b^2 B = 0.$$

Hence we have:

(50) $$\lambda = \pm (b/a')(-B/A)^{1/2}.$$

Equation (49) shows C_1 and C_2 are in involution, as shown also by the projective theory. Equation (50) gives the following theorems:

If the complexes C_1 and C_2 determine a congruence with real directrices, the invariants of these complexes must have opposite signs; and conversely.

By equations (38), (46), and (50) we get the theorems:

If the complexes C_1 and C_2 determine a congruence with real directrices, these complexes will be oppositely twisted; and conversely.

If $\lambda = -(b/a')(-B/A)^{1/2}$, the first special complex of (47) is:

$$
\begin{aligned}
(51)\qquad & a'bD'(\alpha'_1 - \beta'_1 R_1)\omega_{12} - bG^{1/2}(2a'd' + \beta' R_1)\omega_{31} \\
& + a'E^{1/2}(\alpha' + 2bc'R_1)\omega_{23} - 2^8 a'b^2\bar{C}'D'G^{1/2}\omega_{14} \\
& - 2^8 a'^2 b\bar{C}D'E^{1/2}R_1\omega_{24} + a'b(EG)^{1/2}(\alpha'_2 + \beta'_2 R_1)\omega_{34} = 0,
\end{aligned}
$$

where $$R_1 = (-B/A)^{1/2}.$$

To get a simpler form for the axis or first directrix of the congruence determined by C_1 and C_2, let:

$$
(52)\qquad
\begin{aligned}
a'b(\alpha'_1 - \beta'_1 R_1) &= M_1, & bG^{1/2}(2a'd' + \beta' R_1) &= P_1, \\
a'E^{1/2}(\alpha' + 2bc'R_1) &= Q_1, & a'b(\alpha'_2 + \beta'_2 R_1) &= N_1.
\end{aligned}
$$

Then by equation (20), the first directrix is:

$$
\begin{aligned}
(53)\qquad & [\xi - (E^{1/2}G^{1/2}N_1P_1 - 2^8 a'^2 b\bar{C}D'^2E^{1/2}R_1M_1 - FN_1Q_1)/W_1^2]/Q_1 \\
& = [\eta - (E^{1/2}G^{1/2}N_1Q_1 + 2^8 a'b^2\bar{C}'D'^2G^{1/2}M_1 - FN_1P_1)/W_1^2]/-P_1 \\
& = \{\zeta - [2^8 a'b^2\bar{C}'D'G^{1/2}(P_1 - Q_1F/E^{1/2}G^{1/2}) \\
& \quad + 2^8 a'^2 b\bar{C}D'E^{1/2}R_1(Q_1 - P_1F/E^{1/2}G^{1/2})]/W_1^2\}/D'M_1,
\end{aligned}
$$

where $$W_1^2 = Q_1^2 + P_1^2 + D'^2M_1^2 - 2P_1Q_1F/(EG)^{1/2}.$$

This equation shows that when $\alpha'_1 = \beta'_1 R_1$, the first directrix is parallel to the tangent plane or perpendicular to the normal; and that when $P_1 = Q_1 = 0$, it is parallel to the normal.

If $\lambda = (b/a')(-B/A)^{1/2}$, the second special complex of (47) is:

$$
\begin{aligned}
(54)\qquad & a'bD'(\alpha'_1 + \beta'_1 R_1)\omega_{12} - bG^{1/2}(2a'd' - \beta' R_1)\omega_{31} \\
& + a'E^{1/2}(\alpha' - 2bc'R_1)\omega_{23} - 2^8 a'b^2\bar{C}'D'G^{1/2}\omega_{14} \\
& + 2^8 a'^2 b\bar{C}D'E^{1/2}R_1\omega_{24} + a'b(EG)^{1/2}(\alpha'_2 - \beta'_2 R_1)\omega_{34} = 0.
\end{aligned}
$$

To get the form for the axis or second directrix of the congruence determined by C_1 and C_2, let:

$$
(55)\qquad
\begin{aligned}
a'b(\alpha'_1 + \beta'_1 R_1) &= M_2, & bG^{1/2}(2a'd' - \beta' R_1) &= P_2, \\
a'E^{1/2}(\alpha' - 2bc'R_1) &= Q_2, & a'b(\alpha'_2 - \beta'_2 R_1) &= N_2.
\end{aligned}
$$

Then by equation (20), the second directrix is:

$$(56)\quad [\xi - (2^8a'^2b\bar{C}D'^2E^{1/2}R_1M_2 + E^{1/2}G^{1/2}N_2P_2 - FN_2Q_2)/W_2^2]/Q_2$$
$$= [\eta - (E^{1/2}G^{1/2}N_2Q_2 + 2^8a'b^2\bar{C}'D'^2G^{1/2}M_2 - FN_2P_2)/W_2^2]/-P_2$$
$$= \{\zeta - [2^8a'b^2\bar{C}'D'G^{1/2}(P_2 - Q_2F/E^{1/2}G^{1/2})$$
$$- 2^8a'^2b\bar{C}D'E^{1/2}R_1(Q_2 - P_2F/E^{1/2}G^{1/2})]/W_2^2\}/D'M_2,$$

where $$W_2^2 = Q_2^2 + P_2^2 + D'^2M_2^2 - 2P_2Q_2F/E^{1/2}G^{1/2}.$$

We see from this equation that if $\alpha'_1 = -\beta'_1R_1$, the second directrix is perpendicular to the normal; and that if $P_2 = Q_2 = 0$, it is parallel to the normal.

From equations (53) and (56) we deduce the following theorems:

If $P_1 = Q_1 = P_2 = Q_2 = 0$ and if the directrices of the congruence are real, they intersect the tangent plane in two points, such that, when they are joined to the surface point, the two lines thus determined separate the asymptotic tangents harmonically.

These two pairs of lines and the surface normal determine an harmonic pencil of planes with the surface normal as axis.

If $\bar{A} = \bar{B} = 0$, these two points of intersection of the directrices with the tangent plane lie on the directrix of the first kind of the congruence determined by C' and C''. This is easily seen by making $\bar{A} = \bar{B} = 0$ in R_1, α'_1, β'_1, for the two points then become:

$$(57)\quad \begin{aligned} \xi_1 &= -2a'\theta'^{1/2}E^{1/2}/(\theta^{1/2}\alpha_1 - \theta'^{1/2}\beta_1), & \eta_1 &= 2b\theta^{1/2}G^{1/2}/(\theta^{1/2}\alpha_1 - \theta'^{1/2}\beta_1);\\ \xi^2 &= 2a'\theta'^{1/2}E^{1/2}/(\theta^{1/2}\alpha_1 + \theta'^{1/2}\beta_1), & \eta_2 &= 2b\theta^{1/2}G^{1/2}/(\theta^{1/2}\alpha_1 + \theta'^{1/2}\beta_1).\end{aligned}$$

Substitution shows that the points of (57) lie on the line given by (26).

If $\theta = 0$, then $A = 0$ and the complex C_1 is special. Its equation is:

$$(58)\quad 2a'\beta_1D'\omega_{12} + \beta_1\beta_2G^{1/2}\omega_{31} - 2a'c_1E^{1/2}\omega_{23} + 4a'^2D'E^{1/2}\omega_{24} - 2a'\beta_2(EG)^{1/2}\omega_{34} = 0,$$

where $$c_1 = 2c\beta_1 + a'\bar{R}.$$

Its axis in this case, as given by (20), intersects the tangent plane $\zeta = 0$ in the point:

$$(59)\quad \xi = 2a'E^{1/2}/\beta_1, \quad \eta = 0.$$

Its axis can then be written in the form:

$$(60)\quad (\xi - 2a'E^{1/2}/\beta_1)/-2a'c_1E^{1/2} = \eta/\beta_1\beta_2G^{1/2} = \zeta/2a'\beta_1D'.$$

If $\theta' = 0$, then $B = 0$ and the complex C_2 is special. Its equation is:

$$(61)\quad 2b\alpha_1D'\omega_{12} - 2bd_1G^{1/2}\omega_{31} + \alpha_1\alpha_2E^{1/2}\omega_{23} - 4b^2D'G^{1/2}\omega_{14} + 2b\alpha_2(EG)^{1/2}\omega_{34} = 0,$$

where $$d_1 = 2d\alpha_1 + b\bar{R}.$$

Its axis, as given by (20), intersects the tangent plane at:

$$(62) \qquad \xi = 0, \quad \eta = 2bG^{1/2}/\alpha_1.$$

Its axis then takes the form:

$$(63) \qquad \xi/\alpha_1\alpha_2E^{1/2} = (\eta - 2bG^{1/2}/\alpha_1)/-2bd_1G^{1/2} = \zeta/2b\alpha_1D'.$$

Since, by the general theory, the axes intersect when the complexes are both special, by solving equations (60) and (63) for the point of intersection, we get:

$$(64) \qquad \xi = 2a'\alpha_2E^{1/2}K, \quad \eta = 2b\beta_2G^{1/2}K, \quad \zeta = 4a'bD'K,$$

where
$$K = (\alpha_1\beta_2 - 2bc_1)/(\alpha_1\alpha_2\beta_1\beta_2 - 4a'bc_1d_1),$$
and
$$(65) \qquad \alpha_1\beta_2 - 2bc_1 = \alpha_2\beta_1 - 2a'd_1.$$

We get the following theorems from equations (26), (27), (59), (62), and (64).

If the complexes C_1 and C_2 are both special, their axes intersect the asymptotic tangents through the surface point in the same points intercepted by the directrix of the first kind of the congruence determined by C' and C''.

These points of intersection with the asymptotic tangents are at constant distances from the surface point for any angle between the asymptotic tangents.

If the complexes C_1 and C_2 are special, the point of intersection of their axes is always on the directrix of the second kind of the congruence determined by the complexes C' and C''.

If $a'\alpha_2E^{1/2} = b\beta G^{1/2}$, the plane determined by the surface normal and the directrix of the second kind bisects the angle between the planes determined by the surface normal and the asymptotic tangents.

If $\alpha_1 = \beta_1 = K = 0$, the directrix of the first kind becomes the line at infinity in the tangent plane and the axes of the special complexes C_1 and C_2 meet at the surface point.

If $\alpha_2 = \beta_2 = 0$, the directrix of the second kind coincides with the surface normal, the axes of the special complexes C_1 and C_2 intersect on the surface normal at $[0, 0, 2a'D'/(2c\beta_1 + a'\bar{R})]$, while the intersections of the directrix of the first kind with the asymptotic tangents become $(E^{1/2}/2d, 0)$ and $(0, G^{1/2}/2c)$.

Thus we have the following theorem concerning the directrices of the two congruences:

The directrices of the congruence determined by C' and C'', the directrices of the special congruence determined by the special complexes C_1 and C_2, and the asymptotic tangents through the surface point are the six edges of a tetrahedron.

6. *Other congruences determined by the complexes C', C'', C_1, C_2.* The condition (25) for the special complexes of the pencil determined by C_1 and C' is:

$$(66)\qquad (a'\beta'_1 D' + \lambda b D')(-a'\beta'_2 + \lambda b)(EG)^{1/2} + 2^8 a'^2 \bar{C} D' E^{1/2}(\beta' G^{1/2} - 2\lambda b d G^{1/2}) = 0.$$

This reduces to:

$$(67)\qquad \lambda = \pm (a'/b)(-A)^{1/2}.$$

The values of λ for the special complexes of the pencil determined by C_2 and C'' are:

$$(68)\qquad \lambda = \pm (b/a')(B)^{1/2}.$$

From equations (67) and (68) we get the theorems:

If the invariants A and B of C_1 and C_2, respectively, are both positive, then the directrices of the congruence of C_1 and C' are imaginary, while the directrices of the congruence of C_2 and C'' are real.

If the invariants A and B are both negative, then the directrices of the congruence of C_1 and C' are real, while those of the congruence of C_2 and C'' are imaginary.

If the invariants A and B have unlike signs, then either the directrices of both congruences are real or the directrices of both are imaginary.

The values of λ for the special complexes of the pencil determined by C_1 and C'' are:

$$(69)\qquad \lambda = -\bar{A} \pm 2^4 \bar{C} \theta^{1/2}.$$

The values of λ for the special complexes of the pencil determined by C_2 and C' are:

$$(70)\qquad \lambda = -\bar{B} \pm 2^4 \bar{C}' \theta'^{1/2}.$$

From equations (69) and (70) we get the theorems:

If the directrices of the congruence determined by the linear complexes C_1 and C'' are real, the invariant θ must be positive or zero.

If the directrices of the congruence determined by the linear complexes C_2 and C' are real, the invariant θ' must be positive or zero.

We observe that when $\theta = 0$, the value of λ is $-\bar{A}$ and the directrices coincide, as is shown also by the projective theory. Likewise, when $\theta' = 0$, the corresponding value of λ is $-\bar{B}$ and the directrices coincide, which is consistent with the projective theory.

If $\lambda = -\bar{A}$, the common directrix of the congruence of C_1 and C'' is the same as the axis of the special complex C_1.

If $\lambda = -B$, the common directrix of the congruence of C_2 and C' is the same as the axis of the special complex C_2.

7. *Linear congruence determined by the directrices of the congruences of the complexes C', C'', C_1, C_2.* It has been shown by the projective theory that these four complexes have only two lines in common, if θ and θ' are not equal to zero. Then these two common intersectors of the directrices of the congruences of the four complexes become the directrices of a linear congruence. We will investigate this metrically.

By using equations (26) and (27), we found points on these directrices of the congruence of C' and C''. Let any point on the directrix of the first kind be:

$$x_1 = l_1 \cdot 2a'E^{1/2}/\beta_1, \qquad y_1 = m_1 \cdot 2bG^{1/2}/\alpha_1, \qquad z_1 = 0. \tag{71}$$

Let any point on the directrix of the second kind be:

$$x_2 = l_2 \cdot 2a'\alpha_2 E^{1/2}K, \qquad y_2 = l_2 \cdot 2b\beta_2 G^{1/2}K, \qquad z_2 = l_2 \cdot 4a'bD'K. \tag{72}$$

After finding the Plücker line coördinates ω_{ik} for a point of the line joining the points in (71) and (72), we must substitute them in the equations of C_1 and C_2 to determine l_1, m_1, and l_2. We get an unsymmetrical form, unless we assume $\bar{A} = B = 0$.

If θ and θ' are positive, we then have real directrices of the congruence of C_1 and C_2, according to equations (33), (42) and (50). After substituting the ω_{ik} in C_1 and C_2, we get by reduction:

$$\begin{aligned} &2^3a'^2b\bar{C}D'(EG)^{1/2}(-\theta\alpha_1 K \cdot l_1 l_2/m_1 - 2^6\beta_1 KK' \cdot l_2 + 2^6\beta_1) = 0, \\ &2^3a'b^2\bar{C}'D'(EG)^{1/2}(2^6\alpha_1 KK' \cdot l_1 l_2/m_1 + \theta'\beta_1 K \cdot l_2 - 2^6\alpha_1 \cdot l_1/m_1) = 0, \end{aligned} \tag{73}$$

where $\qquad K' = \alpha_1\beta_2 + 2a'd_1 = \alpha_2\beta_1 + 2bc_1$, by (65).

Solving (73), we get two solutions of the quadratic equation:

$$(74) \qquad l_1/m_1 = \pm (\beta_1/\alpha_1)(\theta'/\theta)^{1/2}, \quad l_2 = \pm 2^6/K(\theta^{1/2}\theta'^{1/2} \pm 2^6K').$$

Thus the points of (71) and (72) become:

$$(75) \quad \begin{aligned} &x_1 = \pm 2a'\theta'^{1/2}E^{1/2}/(\theta^{1/2}\alpha_1 \pm \theta'^{1/2}\beta_1), \ y_1 = 2b\theta^{1/2}G^{1/2}/(\theta^{1/2}\alpha_1 \pm \theta'^{1/2}\beta_1), \ z_1 = 0; \\ &x_2 = \pm 2^7a'\alpha_2E^{1/2}/(\theta^{1/2}\theta'^{1/2} \pm 2^6K'), \ y_2 = \pm 2^7b\beta_2G^{1/2}/(\theta^{1/2}\theta'^{1/2} \pm 2^6K'), \\ &z_2 = \pm 2^8a'bD'/(\theta^{1/2}\theta'^{1/2} \pm 2^6K'). \end{aligned}$$

From equations (75) we see that the two lines common to the four complexes, when $\bar{A} = \bar{B} = 0$, intersect the directrix of the first kind in two points which are harmonic conjugates of its intersections with the asymptotic tangents, which agrees with the projective theory.

The investigation of the properties of the cylindroids determined by the axes of systems of two of the four complexes and the ruled surfaces determined by systems of three of the four complexes must be left for a future paper.

UNIVERSITY OF REDLANDS,
REDLANDS, CALIFORNIA.

ON THE REDUCIBILITY OF FAMILIES OF SUBSETS AND RELATED PROPERTIES.

By E. W. Chittenden and Selby Robinson.*

1. *Introduction.* The investigations in abstract sets early revealed the close relationship between the property of Borel, self-compactness, and the closure of decreasing sequences of closed sets.† Analogous properties related to the stronger property of Borel-Lebesgue were found by Sierpinski and Kuratowski ‡ and by R. L. Moore § for the spaces ($\mathfrak{L}$) of Fréchet. These results were later extended to neighborhood spaces in general by Fréchet ¶ and Chittenden ‖ and finally to topological spaces in general by Chittenden.**

These general results together with others of Alexandroff and Urysohn †† suggest propositions which include various forms of the theorem of Lindelöf as well as those of Borel and Borel-Lebesgue as special cases.‡‡

The extension of the theorem of Borel-Lebesgue to general topological spaces effected by Chittenden §§ was based on a theory of coverings of abstract type T. In an article to follow in this journal entitled "Covering theorems in general topology" Robinson has applied this abstract theory in such a variety of ways, that it has been found desirable to develop the subject further. In the form here presented the relation T which entered the original formulation as an abstract form of the relation "interior to", is replaced by the simpler relationship "contains", without loss of generality; since in effect,

* A revision and reformulation of a paper presented to the American Mathematical Society at Columbus, November 27-28, 1931, under the title "On properties of coverings of abstract type related to reducibility."

† See M. Fréchet, *Les Espaces Abstraits*, Gauthier-Villars, Paris (1928), Second Part, pp. 190 ff.

‡ *Fundamenta Mathematicae*, Vol. 2 (1921), pp. 172-178.

§ *Proceedings of the National Academy of Sciences*, Vol. 5 (1919), pp. 206-210.

‖ *Bulletin of the American Mathematical Society*, Vol. 30 (1924), pp. 511-519.

¶ *Bulletin de la Sciences Mathematique*, Series 2, Vol. 42 (1919), pp. 152-156 and *Annales de l'Ecole Normale* (3), Vol. 38 (1921), p. 342.

** *Transactions of the American Mathematical Society*, Vol. 31 (1929), pp. 290-321. This article will be cited frequently as "Topology."

†† *Mathematische Annalen*, Vol. 92 (1924), pp. 258-266.

‡‡ See the excellent report of T. H. Hildebrandt, "The Borel theorem and its generalizations," *Bulletin of the American Mathematical Society*, Vol. 32 (1926), pp. 423-474.

§§ "Topology," p. 306.

if every point of a set E is *interior* to some set V of a family $\mathfrak{V}$, then by replacing each set V by the set W of points of E which are in its interior, we obtain a family of subsets of E whose least common superclass is E. Thus the revised theory reduces to a study of systems $(E, \mathfrak{W}, \mu)$ composed of an arbitrary set E of elements called points, a family $\mathfrak{W}$ of subsets of E, and an *infinite* cardinal number μ. All the theorems stated in "Topology" for coverings of type T follow readily from the results here presented.

The principal problem considered is, if $\mathfrak{W}$, a family of subsets W of E, covers E, that is, if $\Sigma W = E$, where ΣW denotes the class of all points which are elements of some set W of $\mathfrak{W}$, under what conditions does $\mathfrak{W}$ admit a subfamily $\mathfrak{W}_1$ of cardinal number less than μ which also covers E? The conditions obtained require the existence of a subfamily $\mathfrak{W}_0$ of power μ at most which covers E. Satisfactory sets of conditions for the existence of such a family in case E itself is of power greater than μ have not been found.

In sections 2-4 we consider relations among three fundamental properties A, B, C of a system $(E, \mathfrak{W}, \mu)$. In section 5 it is shown by examples that the relations found form a complete set. Section 8 presents a set of three properties similar to A, B, C, which are equivalent for all systems $(E, \mathfrak{W}, \mu)$, and are equivalent to A and B when μ is regular. In sections 10-11 we extend the results of earlier sections to systems $(E, \boldsymbol{W}, M)$, where $\boldsymbol{W}$ is a class of families $\mathfrak{W}$ and M is a class of cardinals μ.

The dual relationship between "contains" and "is contained in" leads to the formulation in section 12 of a number of interesting theorems corresponding to the covering theorems. It is shown that separability is a property of sets intrinsically analogous to the property of Lindelöf. In section 13 we apply the theory of systems $(E, \mathfrak{W}, \mu)$ to sets E in a topological space.

2. *Three fundamental properties.* We begin with the consideration of the following three properties of a system $(E, \mathfrak{W}, \mu)$.*

A. Every subset A of E of power μ determines an element W of the family $\mathfrak{W}$ which contains μ points of A.

B. Every sequence $\mathfrak{S}_\mu$ † determines an element W of $\mathfrak{W}$ which contains a point of every set G_a of $\mathfrak{S}_\mu$.

* These properties correspond to the properties (A), (B), (C), of the theorem at the top of page 302 of "Topology." The statement made there that (B)→(A)→(C) is incorrect. See Theorem 2 below.

† Let $\Omega(\mu)$ denote the least ordinal number such that the class of all smaller ordinals is of power μ. By $\mathfrak{S}_\mu$ we denote a sequence of subsets G_a of E, $0 < a < \Omega(\mu)$, such that for every a, G_a contains all points of G_{a+1} and a point q_a not in G_{a+1}.

C. The family $\mathfrak{W}$ contains a subfamily $\mathfrak{W}_1$ of power less than μ such that any point of ΣW is contained in some set W_1 of $\mathfrak{W}_1$.

It is convenient to introduce the following terminology based on analogies which are justified by the applications. The set A of property A is said to be *nuclear* in $\mathfrak{W}$, while E is *μ-compact.* In property B, the sequence $\mathfrak{S}_\mu$ is *closed* in $\mathfrak{W}$, while E is perfectly μ-compact. In property C, the family $\mathfrak{W}$ is said to be *reducible.* In this terminology, the properties become:

A. E is μ-compact in $\mathfrak{W}$;

B. E is perfectly μ-compact in $\mathfrak{W}$;

C. $\mathfrak{W}$ is reducible to a power less than μ.

These properties depend on μ. The dependence may be indicated by an appropriate subscript when necessary.

3. *A fundamental equivalence.*

THEOREM 1. *If μ is a regular cardinal, $\mathfrak{W}$ is of power μ, and $\Sigma W = E$, properties A, B, and C are equivalent.*

This theorem is an immediate consequence of the following four lemmas.

A set $Q = [q_\alpha / 0 < \alpha < \Omega(/Q/)]$ * is said to be associated with a family $\mathfrak{W}$, if for each element W of $\mathfrak{W}$ there is an index $\beta < \Omega(/Q/)$ such that W does not contain any point of Q of index α greater than β.

LEMMA 1. *A necessary and sufficient condition that property B shall hold, is that there be no subset of E of power μ associated with $\mathfrak{W}$.*

Proof. Suppose that property B holds, but there is a subset $Q = [q_\alpha / 0 < \alpha < \Omega(\mu)]$ of power μ associated with $\mathfrak{W}$. Let $G_\alpha = \sum_{\alpha' > \alpha} q_{\alpha'}$ and let $\mathfrak{S}_\mu = [G_\alpha]$. There is a certain set W of the family $\mathfrak{W}$ which contains a point of each set G_α. But this is impossible as there an index $\beta < \Omega(\mu)$ such that W contains no point q_α of index greater than β, hence no point of G_β. Conversely, if property B does not hold, there is a set of power μ associated with $\mathfrak{W}$. Suppose that $\mathfrak{S}_\mu = [G_\alpha]$ is not closed in $\mathfrak{W}$. For each α let q_α be a point in G_α † but not in $G_{\alpha+1}$.‡ Then $Q = [q_\alpha]$ is associated with $\mathfrak{W}$.

LEMMA 2. *In every system $(E, \mathfrak{W}, \mu)$, $A \rightarrow B$.*

* $/Q/$ denotes the cardinal number of elements in the set Q.

† We assume the axiom of choice.

‡ The set Q so defined is said to be associated with the sequence $\mathfrak{S}_\mu$. This is altogether different from the concept of the set associated with a family $\mathfrak{W}$.

Proof. Let Q be any subset of E of power μ. Property A implies that μ points of Q lie in some set W of $\mathfrak{W}$. Hence Q is not associated with $\mathfrak{W}$. From Lemma 1, property B is present.

LEMMA 3. *If* $/\mathfrak{W}/ = \mu$, $B \rightarrow C$.

Proof. Assume $B\underline{C}$.* Let the elements of $\mathfrak{W}$ be arranged in a sequence $[W_\alpha]$, $0 < \alpha < \Omega(\mu)$. Since $\mathfrak{W}$ is irreducible it contains a subsequence $[W_{\alpha_\beta}]$, $0 < \beta < \Omega(\mu)$, of power μ such that each set W_{α_β} contains a point q_β not contained in any W_α of lower index. Then $Q = [q_\beta]$ is a subset of E of power μ, associated with $\mathfrak{W}$, contrary to Lemma 1.

LEMMA 4. *If* μ *is regular and* $\Sigma W = E$, $C \rightarrow A$.

Proof. Assume $C\underline{A}$. Then E has a subset A of power μ which is not nuclear in $\mathfrak{W}$. But $A < \Sigma W$ and $\mathfrak{W}$ is reducible. Thus the elements of A are contained in less than μ sets W_1. Since $/W_1 \cdot A/ < \mu$ for every W_1 and μ is regular, $\Sigma /W_1 \cdot A/ < \mu$, and we have a contradiction.

4. *Conditions implying the equivalence of the properties A and B.* From Lemma 2, $A \rightarrow B$ in all cases. That $B \rightarrow A$ when μ is regular follows from Lemma 1 and the fact that any subset Q of E of regular power not associated with $\mathfrak{W}$ is nuclear in $\mathfrak{W}$. By modifying an argument due to Sierpinski † we show that B implies A when $/\mathfrak{W}/ = \mu$, and μ is irregular.

Assume $\underline{A}B$. Then there is a subset Q of E, $Q < \Sigma W$, of power μ not nuclear in $\mathfrak{W}$. Let $\mu = L_{\mu_\beta}$, $0 < \beta < \Omega(\nu)$, where ν is the least regular cardinal associated with μ in this manner. Arrange the sets of $\mathfrak{W}$ in a sequence $[W_\alpha]$. For each β, let $T_\beta = Q \cdot \Sigma W_\alpha$, the summation being taken over the sets W_α for which $\alpha < \Omega(\mu_\beta)$ and $/Q \cdot W_\alpha/ < \mu_\beta$. Then $/T_\beta/ \leqq \mu_\beta \cdot \mu_\beta = \mu_\beta$, and $\Sigma T_\beta = Q$. Arrange the points of Q in a sequence $[q_\alpha / 0 < \alpha < \Omega(\mu)]$ in such a way that all points of T_{β_1} precede all points of T_{β_2} if $\beta_1 < \beta_2 < \Omega(\nu)$. Then $Q = [q_\alpha]$ is associated with $\mathfrak{W}$. For any set $W_\alpha \cdot Q$ is contained in some T_β, and therefore contains no point $q_{\alpha'}$ of index α' greater than $\Omega(\mu_\beta)$. By Lemma 1, the existence of the associated set Q contradicts property B. This completes the proof of the following theorem.

THEOREM 2. *In any system* $(E, \mathfrak{W}, \mu)$, $A \rightarrow B$. *If* μ *is regular or* $/\mathfrak{W}/ = \mu$, $A = B$.

---

* The negative of a property P will be represented by $\underline{P}$.

† *Bulletin of the American Mathematical Society*, Vol. 32 (1926), pp. 652-653. It is interesting to observe that the inequality $\mu^\nu > \mu$ follows readily from this argument, as does also the fact that an aggregate of power μ has more than μ distinct subsets of power ν.

5. *Independence examples.* We will show by aid of the following six examples that $A \to B$ is the only relation between the three properties A, B, C, which holds for every system $(E, \mathfrak{W}, \mu)$.

1. $\underline{A}BC$. Let E be the set of all ordinals $\alpha < \aleph_\omega$.* Let $\mathfrak{W}$ consist of all enumerable subsets V of E and all subsets W_n consisting of all ordinals less than $\Omega(\aleph_n)$. Let $\mu = \aleph_\omega$. Evidently A fails, as no set of $\mathfrak{W}$ contains μ points. Any sequence $\mathfrak{S}_\mu$ has an enumerable subsequence running through it. Any enumerable set associated with this enumerable subsequence is a set V of $\mathfrak{W}$ in which $\mathfrak{S}_\mu$ is closed. The family $\mathfrak{W}$ is reducible, since $E = \Sigma W_n$.†

2. $\underline{AB}C$. The system of example 1, with the exception that the family $\mathfrak{W}$ consists only of the sets W_n.

3. $AB\underline{C}$. The system of example 1, with the exception that the family $\mathfrak{W}$ consists only of the sets V, and $\mu = \aleph_0$.

4. $\underline{A}B\underline{C}$. The set E consists of an enumerable family of disjoined subsets E_n, any set E_n being of power $\aleph_n$. Any set E_n consists of a sequence of points e_{n_α} where $0 < \alpha < \Omega(\aleph_n)$. Let $W_{n_\alpha} = \sum\limits_{\alpha' \leqq \alpha} e_{n_{\alpha'}}$ and let the family $\mathfrak{W}$ consist of all sets W_{n_α} and all enumerable subsets V of E. Let $\mu = \aleph_\omega$.

5. $\underline{ABC}$. Let $E = E' + E''$, where E' is the set E used in example 2 and E'' is the set E used in example 4. Let $\mathfrak{W}$ consist of the sets W_n of example 2 and the sets W_{n_α}, together with the sets V, of example 4. Evidently property B fails to hold on E' and C on E''.

6. ABC. Let E be the set of all points of a bounded closed linear interval, the family $\mathfrak{W}$ be the class of all open intervals in E, and $\mu = \aleph_0$.

6. *Properties equivalent to property A.*

THEOREM 3. *In any system $(E, \mathfrak{W}, \mu)$, the following property B is equivalent to property A.*

B. For every sequence $\mathfrak{S}_\mu$ there is a set W of $\mathfrak{W}$ which contains μ elements of each set G_α.

This property $B3$, the property B of theorem 3, implies property $B1$, and is implied by it if μ is regular. To prove $A \to B3$, let $\mathfrak{S}$ be any decreasing sequence of μ subsets G_α of E. Let q_α be an element of $G_\alpha - G_{\alpha+1}$. Then $Q = \Sigma q_\alpha$ is of power μ and must be nuclear in $\mathfrak{W}$ by A. Therefore some W contains μ elements of Q, consequently μ elements of each set G_α. Conversely, if $Q = [q_\alpha]$ is a subset of E of power μ, we may construct a sequence $\mathfrak{S}_\mu$ by

---

* $\aleph_\omega$ is the limit cardinal of the series: $\aleph_1, \aleph_2, \cdots$.

† This example was suggested by one of Sierpinski, *loc. cit.*, p. 650.

the definition $G_\alpha = [q_{\alpha'}/\alpha \leqq \alpha' < \Omega(\mu)]$. The corresponding set W contains μ elements of G_α and therefore of Q.

A subset A of E of power μ is a special case of a family A of μ disjoined subsets of E. Indeed, property A is equivalent to the property:

Every family A of μ disjoined non-null subsets of E determines an element W of $\mathfrak{W}$ which intersects (contains a point of) μ subsets of A.

Properties of this type exist which are equivalent to the respective properties of Theorems 4 and 5 below.

7. *Properties equivalent to property C.* Since property C has been shown to be independent of property A, it is of interest to consider properties analogous to A and B which are equivalent to C. Let E_w denote the set ΣW.

THEOREM 4. *In any system $(E, \mathfrak{W}, \mu)$ the following properties are equivalent.*

A. There is a subfamily $\mathfrak{W}_\rho$ of $\mathfrak{W}$ of power ρ at most, $\rho \leqq \mu$,* in which E_w is ρ-compact.

B. There is a subfamily $\mathfrak{W}_\rho$ of $\mathfrak{W}$ of power ρ at most, $\rho \leqq \mu$, in which E_w is perfectly ρ-compact.

C. The family $\mathfrak{W}$ is reducible to a power less than μ.

Proof. The equality of properties A and B follows from Theorem 2 since E_w is ρ-compact relative to a covering of power ρ. To show that $A \rightarrow C$, let D denote the subset of E_w, necessarily of power less than ρ, not included in ΣW_ρ. Since $E_w - D$ is by hypothesis nuclear in $\mathfrak{W}_\rho$, $\mathfrak{W}_\rho$ is reducible to a power less than ρ by Lemma 3. The required family $\mathfrak{W}_1$ can then be constructed from this reduction of $\mathfrak{W}_\rho$ by the addition of sets of $\mathfrak{W}$ containing elements of D. That $C \rightarrow A$ follows readily from the fact that if $\mathfrak{W}$ is reducible to $\mathfrak{W}_1$ of power $\mu_1 < \mu$, there is a regular cardinal ρ, $\mu_1 \leqq \rho \leqq \mu$, such that every subset of E_w of power ρ is nuclear in $\mathfrak{W}_1$.

8. *Quasi-μ-compactness.* We have shown that the three properties A, B, C of section 2 are equivalent for systems $(E, \mathfrak{W}, \mu)$ for which $\mathfrak{W}$ is of regular power μ and $\Sigma W = E$, and that these properties are not equivalent for all such systems. This raises the question, are there forms of these three properties which are equivalent for all systems of this type which reduce to the fundamental properties A, B, C under the conditions of Theorem 1? The three properties given below fulfill these conditions.

* It is quite easy to show that the cardinal ρ may always be chosen to be regular. If μ is regular and $/\mathfrak{W}/ \geqq \mu$, choose $\rho = \mu$.

It is convenient to make a preliminary definition. A set A is said to be *quasi-nuclear* in $\mathfrak{W}$, if (1) A is nuclear in $\mathfrak{W}$ and $/A/$ is regular, or (2) $/A/$ is irregular and the least upper bound of $/A \cdot W/$ as W varies over $\mathfrak{W}$ is $/A/$.

A. Every subset A of E of power μ is quasi-nuclear in $\mathfrak{W}$, that is, E is quasi-μ-compact in $\mathfrak{W}$.

B. For every sequence $\mathfrak{S}_\mu$ of subsets of E and every regular cardinal $\rho \leqq \mu$, there is a set W in the family $\mathfrak{W}$ and an associated set Q of the sequence $\mathfrak{S}_\mu$ such that $\mathfrak{W}$ contains ρ points of Q, that is, E is quasi-perfectly μ-compact.

C. Every subset C of E which is of power μ contains a subset Q of power μ with respect to which $\mathfrak{W}$ is reducible to a power less than μ.

THEOREM 5. *The three properties stated above are equivalent for all systems* $(E, \mathfrak{W}, \mu)$.

Proof. $A \to B$. Assume that Q is an associated set of a sequence $\mathfrak{S}_\mu$ of subsets of E. Then $/Q/ = \mu$, and consequently for every $\rho \leqq \mu$, some W of $\mathfrak{W}$ contains ρ points of Q.

$B \to C$. Let C be any subset of E which is of power μ. Let $C = [p_\alpha]$, $0 < \alpha < \Omega(\mu)$, and define $\mathfrak{S}_\mu$, by setting $G_\alpha = \sum_{\alpha' \leqq \alpha}^{\Omega(\mu)} p_{\alpha'}$, $0 < \alpha < \Omega(\mu)$. Then C itself is an associated set of $\mathfrak{S}_\mu$. If μ is regular, some set W contains μ elements of C, and $Q = W \cdot C$ is the required set. If μ is irregular, let $\mu = L_{\rho\beta}$, $0 < \beta < \Omega(\nu)$. Then, if Q_β is a subset of C of power ρ_β contained in W_β, we see that $Q = \Sigma Q_\beta$ is of power μ and is contained in $\nu < \mu$ sets W_β of $\mathfrak{W}$.

$C \to A$. Let A be any subset of E of power μ. If μ is regular, and $\mathfrak{W}$ is reducible on a part of A of power μ, then some set W contains μ elements of A as required. If μ is irregular, some part Q of A is contained in less than μ, say μ_1, sets W. If $/Q \cdot W/ \leqq \mu_2 < \mu$, as W varies in $\mathfrak{W}$, we have $/Q/ \leqq \mu_1\mu_2 < \mu$, a contradiction. Hence $/Q \cdot W/$ has the least upper bound μ, as required.

9. *The property* C^*. If $/\mathfrak{W}/ > \mu$, property $B1$ does not imply $C1$. But if the sets W are combined by addition into μ sets W^*, it follows from property $B1$ that the family $[W^*]$ so obtained is reducible. In general, a family $\mathfrak{W}^*$ of sets W^* is said to include a family $\mathfrak{W}$, if each set W of $\mathfrak{W}$ is a subset of some set W^* of $\mathfrak{W}^*$. The following theorem shows the relationship of the properties A, B, and C to the property C^*:

C^*. Any family $\mathfrak{W}^*$ of power μ which includes $\mathfrak{W}$ is reducible to a power less than μ.

The proof of the following theorem is similar to that of Theorem 1 and is omitted.

THEOREM 6. *In any system* $(E, \mathfrak{W}, \mu)$, $A \rightarrow B \rightarrow C^*$ and $C \rightarrow C^*$. *If* μ *is regular and* $\Sigma W = E$, $A = B = C^*$. *If* $/\mathfrak{W}/ = \mu$, $C = C^*$.

Properties C^* and $A5$ hold in all the examples of section 5. We have constructed examples to show that the theorems given present all the relations which hold in general between the five properties A, B, C, $A5$, C^*.

Property B implies that ΣW contains all points of E except those of a subset (which may be null) of power less than μ. Furthermore, if the family $\mathfrak{W}^*$ includes the family $\mathfrak{W}$, ΣW^* has the same property. In view of this fact, it is evident that the set ΣW plays a more important role than E, and we shall lose nothing by assuming henceforth that $\Sigma W = E$.

10. *Systems* $(E, \boldsymbol{W}, M)$. The classical forms of the theorem of Borel relate to families of coverings, and classes of cardinal numbers. The theorems stated for a single covering and cardinal number in the preceding sections readily supply groups of equivalent properties for systems of the type $(E, \boldsymbol{W}, M)$, where $\boldsymbol{W}$ denotes a class of families $\mathfrak{W}$, and M a class of cardinals μ. Of the many theorems possible the following seem to be of the greatest interest and value.

Consider the following three properties.

A. Every subset of E whose power is in the class M, is nuclear in every family $\mathfrak{W}$ of the class $\boldsymbol{W}$.

B. Every sequence $\mathfrak{S}_\mu$, where μ is in the class M, is closed in every family of $\mathfrak{W}$.

C^*. Every family W^* whose power is in the class M and which includes a family of the class $\boldsymbol{W}$ is reducible.

THEOREM 7. *For every system* $(E, \boldsymbol{W}, M)$, $A \rightarrow B \rightarrow C^*$. *If every cardinal in* M *is regular and* $\Sigma W = E$ *the three properties are equivalent.*

If ν is the regular cardinal determined by an irregular cardinal μ, closure in a family $\mathfrak{W}$ of all sequences $\mathfrak{S}_\nu$ implies the closure of every sequence $\mathfrak{S}_\mu$ in $\mathfrak{W}$. From this fact and Theorem 7 we easily derive the following result.

THEOREM 8. *If* M *is the class of all infinite cardinals less than some fixed cardinal* $\bar{\mu}$, $A8 = B7 = C^*7$, *where* $A8$ *is defined as follows.*

*A*8. All infinite subsets of E of regular power less than $\bar{\mu}$ are nuclear in every family $\mathfrak{W}$ of the class $\boldsymbol{W}$.

11. *Conditions for the reducibility of all coverings of a given class.* In preceding sections we have observed the effect of the assumption $/\mathfrak{W}/ = \mu$. Consider the following three properties.

A. Any family $\mathfrak{W}$ of a class $\boldsymbol{W}$ has a subfamily $\mathfrak{W}_1$ in which E is $/\mathfrak{W}_1/$-compact.

B. Any family $\mathfrak{W}$ of a class $\boldsymbol{W}$ has a subfamily $\mathfrak{W}_1$ in which E is perfectly $/\mathfrak{W}_1/$-compact.

C. Every family of $\boldsymbol{W}$ is reducible.

THEOREM 9. *In any system* $(E, \boldsymbol{W})$, $A = B = C$.

Proof. Apply Theorem 4 with $\mu = /\mathfrak{W}/$.

Another equivalence is secured by an application of Theorem 4 to all families of $\boldsymbol{W}$ with $\mu = \lambda$, the least cardinal in M. The reducibility property thus secured is the property

$C10$. Each family of the class $\boldsymbol{W}$ is reducible to a power less than λ.

THEOREM 10. *Let* $\lambda = \aleph_0$. *Suppose that a family* $\mathfrak{W}_1$ *belongs to the class* $\boldsymbol{W}$ *if* $\mathfrak{W}_1$ *covers* E *and is a subfamily of a family of* $\boldsymbol{W}$. *Then property* $C10$ *is equivalent to each of the properties of Theorems 7 and 9.**

This theorem follows from the following three lemmas which can be established quite easily by the aid of Lemmas 2-4 and Theorem 6.

LEMMA 5. *In any system* $(E, \boldsymbol{W}, M)$, $C10 \rightarrow C^*7 \rightarrow C9$.

LEMMA 6. *In any system* $(E, \boldsymbol{W}, M)$ *for which the subfamilies* $\mathfrak{W}_1$ *of families* $\mathfrak{W}$ *of* $\boldsymbol{W}$ *belong to* $\mathfrak{W}$ *whenever they cover* E, $C10 = C^*7 = C9$.

LEMMA 7. *In any system* $(E, \boldsymbol{W}, M)$ for which $\lambda = \aleph_0$, $C10 \rightarrow A7$.

12. *Duality.* If in property $A1$ the family $\mathfrak{W}$ and the set E are interchanged and also the relations: "contains" and "is contained in", the following property results.

A. For every subfamily $\mathfrak{W}_\mu$ of power μ of sets of the family $\mathfrak{W} = [W]$ there is a point of E which is contained in μ sets of $\mathfrak{W}_\mu$.

Assuming that no set W is null † property $C1$ corresponds to:

C. There is a subset Q of E of power less than μ such that each set W of $\mathfrak{W}$ contains a point of Q.

Then the following theorem is established by the same arguments as Theorem 1, and is in fact a dual of part of that theorem.

---

* The theorem on page 306 of "Topology" follows readily from the equivalence just obtained between the properties $C10$, $A7$, and $B7$. Cf. Theorem 13 below.

† We also assume that $\Sigma W = E$.

THEOREM 11. *If μ is regular and $/E/=\mu$, properties A and C are equivalent.*

An application of the same procedure to Theorem 4 yields

THEOREM 12. *In any system $(E, \mathfrak{W}, \mu)$ a necessary and sufficient condition for property C11, is that there be a subset E_ρ of E of power $\rho \leqq \mu$ such that for any subfamily $\mathfrak{W}_\rho$ of $\mathfrak{W}$ of power ρ there is a point of E which is contained in ρ sets of W_ρ.**

13. *An application to topological spaces.* If we assume a system (P, K) as defined in "Topology", we may consider systems $(E, \mathfrak{V}, \mu)$, where $\mathfrak{V}$ denotes a family of neighborhoods V of points of E. If W denotes the subset of E which is interior to V, and $\Sigma W > E$, where the summation is taken over the entire family $\mathfrak{V}$, then $\mathfrak{V}$ is a *proper covering* of E.

A necessary and sufficient condition that E be properly μ-compact in itself in the space P is that E be properly μ-compact in every one of its proper coverings $\mathfrak{V}$.

Proof. If E is not properly μ-compact in itself it admits a proper covering in which it is not properly μ-compact. Conversely, if E admits a proper covering in which E is not properly μ-compact, there is a subset A of E of power μ which has no point of E as a proper μ-point.

From this proposition and Theorem 10 we obtain a result from "Topology" already cited.

THEOREM 13. *A necessary and sufficient condition that every infinite proper covering of a set E be reducible is that every infinite subset of E have a proper nuclear point in E.*

For further applications the reader is referred to the article by Robinson.

* Property C11 is related to separability. Corollaries of Theorems 11 and 12 provide necessary and sufficient conditions for separability as is shown by Robinson in the article previously mentioned.

A CHARACTERISATION OF THE CLOSED 2-CELL.*

By Leo Zippin.

In this paper we establish the following characterisation of a closed 2-cell (a set of points homeomorphic with a plane circle plus its interior):

If C is a continuous curve (compact)* containing a simple closed curve J and at least one arc ab such that ab has its endpoints and these only on J,† *and such that* every arc spanning † J irreducibly separates ‡ C, then C is a closed 2-cell (and J, of course, its boundary).

We shall set about the proof in this fashion that we shall show that $C - J$ is topologically a euclidean plane (therefore homeomorphic to the *interior* of a plane circle) and has every point of J for limit point. We shall regard this as by far the major part of the argument, for we shall then know essentially that our point set is an open non-singular 2-cell bounded by a simple closed curve. At that point it is still conceivable that we have a singular 2-cell, the singularities being on the boundary. But it is as least intuitively clear that the condition that spanning arcs must separate can be invoked to rule out this situation. However, the methods of the paper do not lend to easy translation into combinatorial technique and we shall have to close the proof by carrying out a "mapping" of C onto a closed 2-cell.

It is obvious from the nature of our conditions that there is little we can say about C, in the beginning at least, excepting through the point set J. The first thing we shall have to determine, and in the proof of this lies most of the novelty of this paper, is this that every spanning arc of C separates C into two components each of which contains a point of J. This is a consequence of the assertion which we now prove that *if xy is a spanning arc and A any component of $C - xy$, A contains a point of J.* For $J - (x + y)$ is

* Throughout this paper continuous curves shall be supposed merely locally compact except when explicitly restricted; as in this theorem. For definitions, etc., one may be referred to an earlier paper of the author: "On continuous curves and the Jordan curve theorem," *American Journal of Mathematics*, Vol. 52 (1930), pp. 331-350. We refer to it as J. C., and shall need one of its principal results.

† In symbols, $ab \cdot J = a + b$; we shall say that such an arc "spans" J, or is a "spanning arc."

‡ A subset K irreducibly separates C if $C - K$ is not connected, but $C - K''$ is connected whenever K'' is a *proper* subset of K. The reader will know that this implies for closed subsets K of a continuous curve C that every component of $C - K$ has every point of K for limit point.

the sum of two open arcs and each can belong to at most one component of $C - xy$. We shall achieve the proof by a reductio ad absurdam.

1. *We suppose then that there does exist in C a spanning arc xy such that in $C - xy$ there is a component A which contains no point of J.*

Since the argument is not to be very short, it may well be anticipated here, informally. We are going to show that A really must have something that corresponds to an outer edge, a sort of boundary. We shall show that this *edge* must be connected and have x and y as limit points. We shall construct it so that it does belong to $\bar{A}$ (and its construction will depend on the compactness of $\bar{A}$) but has no point in A or on the open arc xy. It will follow that it must belong to J, has in fact to coincide with one of the arcs of $J - (x + y)$. This will not prove that *every* point of this arc is a limit point of A because the argument is based on the assumption that no point of J is a limit point of A (excepting x and y). It will contradict this assumption. See note to 1. 5.

1. 1. $C - xy$ certainly contains at least one component B which contains a point of J. Of course, $\bar{B}$ contains xy, since this is true of every component. Then if sq is any subarc of $\langle xy \rangle$ * there exists in $\bar{B}$ an arc fg such that $\langle fg \rangle \subset B$, $fg \cdot J = f$, $fg \cdot xy = fg \cdot (\langle sq \rangle) = g$. *For* $B + \langle sq \rangle = \bar{B} - (xy - \langle sq \rangle)$ is an open subset of the continuous curve † $\bar{B}$ and is connected because B is connected and $\bar{B}$ contains sq. Therefore $B + \langle sq \rangle$ is a continuous curve (it is obviously locally compact) and, of course, arcwise connected. In particular then it contains an arc X joining some point of the open arc sq to some point of that open arc of J which is contained in B. But the closures of these last two arcs are mutually exclusive, since $sq \subset \langle xy \rangle$. and $wy \cdot J = x + y$. Then the arc X certainly contains at least one subarc which has one and only one point on each of the open arcs above, and this is the desired arc fg.

1. 2. We can now show that there do not exist in $\bar{A}$ two mutually exclusive arcs pq and st such that each (excepting for its endpoints) belongs to A, and these endpoints are distinct inner points of xy in order: $xpsqty$. For let fg be the arc in 1. 1 above. Then the arc $xpqgf$, where xp and qg belong to xy, spans J but does not separate C. *For* every component of $C - xpqgf$ must contain points of A since it has points of A (namely points of $\langle pq \rangle$) for limit points, and A is open. And it must also contain points of B since

* Read: the open arc xy.

† It is well-known that if K is a subcontinuous curve of a continuous curve C, and D any component of $C - K$, then $K + D$ is a continuous curve.

it has points of $\langle fg \rangle$ for limit points. Then, being connected and containing points of A and B it must contain points of xy. But it is clear that all points of xy not on the arc $xpqgf$ lie in a connected subset of $C - xpqgf$, that one namely which contains st. Then there is only one component, and $xpqgf$ does not separate. This is a contradiction.

1. 3. If xzy is any arc of $\bar{A}$ it must span J, since $\bar{A} \cdot J = (x + y)$, and must irreducibly separate C. If it does not coincide with our arc xy, we have in mind the fixed arc xy which determines the component A, it must contain a point of A; for $\bar{A} = A + xy$. In this case every component of $C - xzy$ contains points of A, and if each of them contained points not in A each of them would necessarily contain points of the original arc xy. But the points of xy not on xzy belong to a single component of $C - xzy$, for example the one which contains the connected set B (of 1. 1). It follows that $C - xzy$ contains one and only one component, which we denote by A_z such that A_z belongs to A. This component obviously contains no point of J. Now by an entirely similar argument, if $xz'y$ is any arc of $\bar{A}_z$, $C - xz'y$ has a single component $A_{z'}$ which belongs to A_z. While not necessary to us it will make things a little more vivid to notice that these components A_z and $A_{z'}$ are also singled out as the *only* components of $C - xzy$ and $C - xz'y$ respectively which contain no points of J. For it follows by a slight addition to the argument of 1. 2 that only one such component exists for a given arc. This is a consequence of the easily noticed fact that it is not essential to the argument of 1. 2 that pq and st belong to the same component A but only that neither of them belongs to B.

1. 4. We wish to show that if z is any point of A there exists an arc szt, $\langle szt \rangle \subset A$, $s + t \subset \langle xy \rangle$. Now $A + \langle xy \rangle$ is a continuous curve, A is connected and $\bar{A} \supset \langle xy \rangle$. It follows readily (Whyburn) that $A + \langle xy \rangle$ contains a maximal cyclicly connected subset M * which contains $\langle xy \rangle$. If M does not coincide with $A + \langle xy \rangle$ it contains a point z' which is a cutpoint of the latter set, and it contains an arc $s'z't'$ such that $s'z't' \cdot (\langle xy \rangle) = s' + t'$ (Ayres) †: it is obvious of course that the point z' could not have been a point of xy. But the arc $xs'z't'y$ determines a component $A_{z'}$ contained in A; $\bar{A}_{z'} \supset xs'z't'y$. Then $(A + \langle xy \rangle) - z'$ contains a component which contains every point of $\langle xy \rangle$. Then if z' separated some two points of $A + \langle xy \rangle$ it would have to separate some point of this set from the arc $\langle xy \rangle$. But

* M is, of course, a continuous curve. We must assume acquaintance with the ideas of cyclic-connectivity. One is referred to the paper by Whyburn and Kuratowski in *Fundamenta Mathematicae*, Vol. 15 (1930).

† See also note to 2. 2.

$x + z' + y$ as proper subset of a spanning arc $xs'z't'y$ does not separate C. Therefore $C - (x + z' + y)$ contains an arc joining any given point of $A - z'$ to some point of $\langle xy \rangle$, and this arc has a subarc, certainly, which belongs to $A + \langle xy \rangle$. The contradiction shows that $A + \langle xy \rangle$ coincides with M, and is therefore cyclicly connected. Then it follows again that if z is any point of A, there is an arc szt, $\langle szt \rangle \subset A$, $s + t \subset \langle xy \rangle$. If A_z is the component of $C - xzy$ contained in A, it follows by entirely similar argument that $A_z + \langle xszty \rangle$ is cyclicly connected.

1. 5. Actually we shall require less than the last remark above. It is sufficient for us that z does not separate $A_z + \langle xzy \rangle$. For in this case there certainly exists an arc $s''z''t''$, say, such that $\langle s''z''t'' \rangle \subset A_z$, $s'' \subset \langle xz \rangle$ and $t'' \subset \langle zy \rangle$. The arc $xs''z''t''y$, let us call it $xz''y$, determines a component $A_{z''}$ contained in A_z. Clearly z belongs to $C - \bar{A}_{z''}$. We shall say that the arc $xz''y$ *covers* the point z.* It is important to note that this covering arc is of a certain simple type relative to its intersection with the arc xy. Specifically the intersection of xy with $xz''y$ coincides in some neighborhood of $x + y$ with the arc xy itself. We shall say that such an arc is of *admissable type*. The relevance of this will lie in the fact that any arc which is contained in the sum of two arcs of admissable type is also of admissable type. It is to be noted that we have not required that an arc of admissable type intersect xy in *precisely* two arcs.

1. 6. We have shown that if z is any point of $\bar{A} - (x + y)$ there exists an arc $xz'y$ of admissable type covering z, i. e. such that z belongs to $C - \bar{A}_{z'}$. From the separability of A it follows that there exists a *countable* set of arcs, xz_1y, $xz_2y, \cdots$, such that if z is any point of $\bar{A} - (x + y) = A + \langle xy \rangle$, there exists an integer n depending on z such that $z \subset C - \bar{A}_{z_n}$, where A_{z_n} is that component of $C - xz_ny$ which is contained in A. Now if the continuous curves $\bar{A}_{z_n}$, $n = 1, 2, \cdots$, formed a monotonic decreasing set we could conclude at once that there was a continuum K common to all of them, and that this continuum contains x and y. It is this continuum K which we are seeking. To secure it we shall be obliged to find another sequence of continuous curves, essentially equivalent to the first, but in addition monotonic decreasing. To achieve this we shall have to show that *if xzy and $xz'y$ are two arcs of admissable type with A_z and $A_{z'}$ as the corresponding components, then there exists an arc $xz''y$ of admissable type with a component $A_{z''}$ contained in $A_z \cdot A_{z'}$.*

*** It is at this point that we enter in earnest on the construction of the "edge" whose existence we anticipated in 1.**

1.7. We must first verify that if no point of $xz'y$ belongs to A_z, A_z is contained in $A_{z'}$. If $A_z \not\subset A_{z'}$, A_z must contain a point not in $A_{z'}$. If it also contains a point of $A_{z'}$ it must contain a point of $xz'y$. Since A_z contains no point of $xz'y$, we have that $A_z \cdot A_{z'} = 0$. But now, because each of our arcs is of admissable type, it follows that there is an arc x' of our original arc xy which belongs to xzy and to $xz'y$. There is no difficulty in showing that there exists in $\bar{A}_z$ an arc pq such that p and q are inner points of x' and $\langle pq \rangle \subset A_z$ and that there exists in $\bar{A}_{z'}$ an arc st such that s and t are inner points of x' and $\langle st \rangle \subset A_{z'}$ *and the order* of points on xx' is: $xpsqtx'$. But this contradicts 1.2. Then in this case $A_z \subset A_{z'}$ and xzy is the desired arc.

1.8. Then we may suppose that there exist points of $xz'y$ which belong to A_z. If t is such a point there corresponds to it an arc $t'tt''$ of $xz'y$, with endpoints on xzy and belonging except for its endpoints to A_z. Let us call all such arcs of $xz'y$ *relevant* arcs: every point of $xz'y \cdot A_z$ belongs to one such relevant arc, and any two of these arcs have at most one endpoint in common. We see that the set of these arcs is countable, and that the set of their diameters converges to zero. Each relevant arc $t'tt''$ determines uniquely a certain *major* arc (also a relevant arc): we shall say that a relevant arc $s'ss''$ is a major arc provided that if $t'tt''$ is any other relevant arc then the subarc $s's''$ of xzy is not contained in the subarc $t't''$ of xzy.* Now if $s'ss''$ and $t'tt''$ are two major arcs, we see that their endpoints on xzy cannot overlap (this is the essence of 1.2). It follows readily that there exists in xzy plus the set of major arcs of $xz'y$ an arc $xz''y$ *which contains the set of all major arcs.*

Now it follows at once that if $A_{z''}$ is the component of $C - xz''y$ contained in A, $A_{z''}$ is contained in A_z: for every point of $xz''y$ belongs to $\bar{A}_z$. We want to know that $A_{z''}$ also belongs to $A_{z'}$. This shown, our argument is completed. For this we need merely to prove that no point of $xz'y$ belongs to $A_{z''}$ (1.7). Now if t is a point of $xz'y$ which does belong to $A_{z''}$, t is also a point of A_z and belongs to what we have called a relevant arc $t'tt''$ of $xz'y$. But this cannot be a major arc, since all of these belong to $xz''y$. Then it follows that there exists an arc $s'ss''$, which is a major arc, such that the arc $s's''$ of xzy includes the arc $t't''$ of xzy. Clearly either t' or t'', say t', must be an inner point of $s's''$. Then the arc tt' has no point in common with the arc $xz''y$. But t' is a point of xzy. Then if t is a point of $A_{z''}$, it follows that t', also, is a point of $A_{z''}$, and this we know to be impossible. Then, finally, $A_{z''} \subset A_z \cdot A_{z'}$, and $xz''y$ is the desired arc.

---

* Our entire argument will seem very familiar to those acquainted with R. L. Moore's early "Foundations of plane analysis situs." It occurs in a very similar connection in one of his theorems there.

1. 9. We come now to the final argument of this section which will contradict the assumption in 1. By the preceding arguments there exists an arc, call it $x(1)y$, belonging to $xz_1y + xz_2y$ (therefore to $\bar{A}$) such that the component, call it A_1, of $C - x(1)y$ belonging to A is contained in $A_{z_1} \cdot A_{z_2}$. By a repetition of the argument there exists an arc $x(2)y$ such that the corresponding component A_2 is contained in $A_1 \cdot A_{z_3}$. Continuing inductively we see that there exists a sequence of arcs $x(1)y, x(2)y, \cdots, x(n)y, \cdots$, and corresponding components $A_1, \cdots, A_n, \cdots$, such that *a*) the arcs and components are contained in $\bar{A}$, *b*) A_{n+1} is contained in A_n and this in turn is contained in $\prod_{i=1}^{n} A_{z_i}$. Now, as a monotonic decreasing sequence of continua in a compact space $\bar{A}$, the $\bar{A}_i$ have a common part which is a continuum K, and this contains the points x and y which belong to every $\bar{A}_i$. We have that $K \subset \bar{A} = A + \langle xy \rangle + (x + y)$. But if z is any point of $A + \langle xy \rangle$, for some n, z fails to belong to $\bar{A}_{z_n}$, therefore to $\bar{A}_n$, therefore finally to K. Then K is a subset of the point-set $x + y$. But K is a continuum and contains x and y. This is certainly impossible.

Therefore, finally, the assumption of 1 is untenable, and we may assert that if xy is any spanning arc of C, $C - xy$ contains two components, each of these containing one of the open arcs of $J - (x + y)$.

2. Let M be the component of $C - J$ containing $\langle ab \rangle$, and let $C' = M + J$. If we can show that C' is a closed 2-cell, it will follow at once that C must coincide with C' and is therefore a 2-cell: the argument is trivial. Let us verify that C' has all of the properties of C with this additional one that J does not separate it. The last, of course, follows from definition of C'. Clearly C' is a compact continuous curve, contains J and a spanning arc. And clearly any spanning arc xy separates C' between points of J, because $C' \subset C$. *Suppose now* that γ is a subarc of $\langle xy \rangle$ such that $xy - \langle \gamma \rangle$ separates C' between some pair of points p and q. There is an arc pq in $C - (xy - \langle \gamma \rangle)$. If $pq \cdot J = 0$, $pq \subset M \subset C'$, contrary to supposition. Let p' and q' be the first points of pq on J, from p and q respectively. If they belong to the same arc of $J - (x + y)$, pp' (of pq) $+ p'q'$ (of J) $+ q'q$ (of pq) belongs to C'; again contrary to supposition. Therefore p' and q' belong to different arcs of J and $p'q'$ (of pq) must contain a point of $\langle \gamma \rangle$. Let p^* and q^* be the first points of $p'q'$ on γ, in order from p' and q' respectively, and let p'' and q'' be the first points of p^*p' and q^*q' respectively (in the order written) which are on J. There is no difficulty in finding, in a sum of the arcs above, an arc pq in $C' - (xy - \langle \gamma \rangle)$. Then our supposition is untenable, and C' has the property that *every spanning arc separates it*

irreducibly. From the first part of this paper, every spanning arc separates C' into two components each of which contains an arc of J. We know that every point of $C - J$ belongs to some spanning arc. In particular then every point of M belongs to a spanning arc in C', and does not separate C' from what we have seen above. Then C' is cyclicly connected. For the next moments our main concern is with M. We wish to show that M is homeomorphic with the ordinary euclidean plane.

2. 1. Let us verify that M contains some simple closed curve. Let α denote any subarc of $\langle ab \rangle$. C' contains an infinity of arcs $\{\beta\}$ each having an endpoint on α, an endpoint on J, and no other point on $J + ab$, the endpoints on α being distinct: clearly each of these arcs belongs to M, excepting for its endpoint on J. If any two of the open arcs of the set $\{\beta\}$ have a common point, the desired simple closed curve exists in their sum with α. But otherwise, remembering that M is compact in a suitable neighborhood of α, we can extract from the set $\{\beta\}$ a sequence of arcs with endpoints on α which converge to some limiting continuum. Using the local connectedness of M, the desired simple closed curve is easily constructed in the sum of α, some two of the arcs of this sequence, and a fourth connecting them at some slight remove from α. Then M contains at least one simple closed curve.

2. 2. Let us verify that every simple closed curve K of M separates M. Since C' is cyclicly connected, $K \cdot J = 0$, and both sets are closed, there exist (Ayres *) two mutually exclusive arcs a_1b_1 and a_2b_2 such that $\Sigma\, a_i \subset J$, $\Sigma\, b_i \subset K$, $(\Sigma \langle a_ib_i \rangle) \cdot (J + K) = 0$. For future use, it will not hinder our present argument, let γ denote an arbitrary arc of K, and let b_1b_2 denote one of the arcs of K which contains a point not on γ. Let A denote that component of $C' - a_1b_1b_2a_2$ which does not contain the other arc, with endpoints b_1 and b_2, of K. There exists an arc a_3b_3, $\langle a_3b_3 \rangle \subset A$, $a_3 \subset J$; $b_3 \subset \langle b_1b_2 \rangle$ and is not a point of γ. In that component of $C' - a_1b_1b_3a_3$ which does not contain b_2, there is an arc $\langle a_4b_4 \rangle$, $a_4 \subset \langle a_3a_1 \rangle$ (of J), $b_4 \subset \langle b_3b_1 \rangle$ (of K), and no point of b_3b_4 is a point of γ. We shall have no further use for γ than this observation that *every arc of K belongs to a spanning arc of C':* in the case above, γ belongs to $a_4b_4b_1b_2b_3a_3$.

We shall need the configuration: $K + J + \sum_{1}^{4} a_ib_i$. We observe that it was constructed to have these properties: 1) $\Sigma\, a_i \subset J$, in order 1234, 2) $\Sigma\, b_i$

* This is also a special case of a final corollary in a paper (of ours): "Independent arcs of a continuous curve" to appear in the *Annals of Mathematics*, January, 1933.

$\subset K$, in order 1234, 3) $(K+J)\cdot(\Sigma\langle a_ib_i\rangle)=0$, 4) the arcs a_ib_i are mutually exclusive. Now let B denote that component of $C'-a_4b_4b_1b_2b_3a_3$ which does not contain a_1+a_2. There is in B an arc $\langle kk''\rangle$, where k'' lies on the arc $\langle a_3a_4\rangle$, not containing a_1+a_2, of J, and k correspondingly on the arc b_1b_2, not containing b_3+b_4 of K. Then, in order from k, the arc kk'' has a first point k' on the arc b_3b_4 of K: otherwise the spanning arc $a_4b_4b_3a_3$ does not separate C' between points of J. Now it is easy to see that K separates M between every pair of points such that one of them is on $\langle kk'\rangle$ and the other on some arc $\langle a_ib_i\rangle$. Then we may conclude that every simple closed curve of M separates M. It remains to verify that no arc of a simple closed curve of M separates M.*

2.2. We have seen above that every such arc belongs to a spanning arc. By a trivial argument we can reduce our problem to showing this: that if xy is any spanning arc of C' and z any point of $\langle xy\rangle\subset M$, then every point of $M-xz$ can be joined in that set to a point of $\langle zy\rangle$. There is no difficulty in showing that every point of $M-xz$ can be joined in that set to a point of an arc mm', where m' is on one of the open arcs of J and m is on $\langle xzy\rangle$, and the arc mm' has its endpoints only on $J+xzy$. If we can always find this arc mm' such that $m\subset\langle zy\rangle$, our argument is concluded: suppose, then, that for some point of $M-xz$, the corresponding arc mm' has m on $\langle xz\rangle$. Then if we can find an arc with an endpoint on $\langle mm'\rangle$ and an endpoint on $\langle zy\rangle$ and these points only on $J+xzy+mm'$, we are through again. Suppose then that no such arc is to be found. But, still further, if we can now find an arc $m'm''$, where m'' is on $\langle zy\rangle$, and $m'm''\cdot(J+xzy+mm')=m'+m''$, our argument is at an end. *For* in that component, say Z, of $C'-m'mzy$ which does not contain x there is an arc joining a point of $m'y$ (of J) to a point, say p, of $\langle mm'\rangle$. Since $m'm''y$ is a spanning arc, the arc above must intersect it and has, accordingly, a first point on it, in order from p: call the point q. Then $pq+qm''$ is the arc we have sought.

But now to show the existence of an arc $m'm''$, above, under the suppositions of the paragraph above that certain other arcs do not exist reduces to a curious sort of accessibility argument which we have given once under circumstances so closely analogous,† that we shall dispense with it here. We consider it proven, then, that no arc of a simple closed curve of M separates M, that every simple closed curve of M does separate M, and that M contains

* J. C., Theorem 3″, page 341 (in the light of the second of the two small paragraphs immediately following it).

† J. C., pp. 343-5, § 4. 3.

at least one simple closed curve. Then we know that M is homeomorphic to the ordinary euclidean plane.*

3. We are in a position to introduce in M any convenient coördinate system. Thus, let aob (o is a point of "reference") be any spanning arc of C' ($\langle aob \rangle \subset M$), let z be any point on one of the open arcs of J, let A be that component of $C' - aob$ which contains z, and let A^* be that component of $M - \langle aob \rangle$ which lies in A. Then we may suppose that M has been so "ruled" that o is the origin of coördinates, the ray oa is the positive *x-axis,* the ray ob the positive *y-axis,* and A' the first quadrant. It should not confuse the reader that we now denote numerical coördinates by (x, y) although we were accustomed earlier to use these symbols for points of C. Now let a_n be the point $(n, 0)$, b_n the point $(0, n)$ and a_nb_n the "quarter-circle" with "center" at the origin o, $n = 1, 2, \cdots$, and finally let A_n denote the set of points of A' exterior to the "circle" of radius n. Now the sets $\bar{A}_n$ (of C') form a monotonic decreasing sequence of continua. Therefore their common part, call it K, is a continuum, and it contains a and b. Now $K \subset \bar{A}' \subset \bar{A} = aob + A' + azb$, the last being an arc of J. But it should be clear that for every point of $\langle aob \rangle + A'$ there exists an integer n such that this point is not contained in $\bar{A}_n$. Therefore K is a subset of azb, and being connected and containing a and b it must coincide with azb.† Then, in particular, the point z belongs to $\bar{M} \supset \bar{A}'$. But z was a quite arbitrary point of J. Therefore every point of J is a limit point of M.

We observe, also, that the arcs aa_nb_nb converge to azb: i. e. if Z denotes an arbitrary neighborhood of azb, at most a finite number of the arcs of the sequence can have points exterior to Z. *For,* otherwise, there exists a point k, not on azb, and an infinite set of points drawn from *distinct* arcs aa_nb_nb, therefore from distinct sets $\bar{A}_n$, converging to k. But this implies, by a quite trivial argument, since the A_n's form a monotonic sequence, that k belongs to every $\bar{A}_n$, therefore to their common part K. This is a contradiction. Then we have shown also that if azb is any arc of J and there exists a spanning arc aob, then there exists a spanning arc ab whose diameter differs by arbitrarily little from that of azb. But since every point z of J is a limit point of M, it follows in turn that there exists a spanning arc ab, $a \neq z \neq b$, such that it and the arc azb of J are of arbitrarily small diameter. Again, since z is not a cutpoint of C' (we have shown that C' is cyclicly connected) it

* It is clear, of course, that M is not compact, but that every simple closed curve determines in it one compact domain.

† The parallel with the first part of this paper is worth remarking. In the first part we were led to contradiction for want of a boundary arc.

follows without difficulty that given any $\epsilon > 0$ we can choose a δ, $\epsilon > \delta > 0$, such that for every spanning arc ab, above, such that it and the arc azb are of diameter less than δ, the component A_z of $C' - ab$ which contains z is of diameter less than ϵ.* It now follows, by familiar arguments, that every point of J is arcwise accessible from M. Now it should be clear that if ab is any spanning arc of C', z a point on one of the arcs of J, A the component of $C' - ab$ containing z, and A' the component of $M - ab$ contained in A, *that* $\bar{A} + ab$ has all of the properties, and their consequences above, of C' with $abza$ replacing the simple closed curve J and A' replacing M. It follows that if we are given any finite set of points $a_1, \cdots, a_n$ on J in order as written, there exists a set of n arcs $a_i a_{i+1}$ (mod n) such that the corresponding open arcs are mutually exclusive and contained in M. Further, if P denotes the "polygon" whose "edges" are the arcs above, and D denotes that component of $C' - P$ which contains no point of J, then it can be shown without difficulty that $\bar{D}$ is a 2-cell with boundary P.

4. Since J is homeomorphic with a circle it is clear that we might have supposed the integer n, above, sufficiently large, and the points a_i uniformly distributed around J (in the sense of the homeomorphism carrying J into a circle) so that the arcs $a_i a_{i+1}$ *on* J are arbitrarily small. Then in this case, from what we have shown above, the arcs $a_i a_{i+1}$ *of* M may be supposed arbitrarily small and it follows further that if D_i denotes the component of $M - a_i a_{i+1}$ which is bounded by the arc $a_i a_{i+1}$ of J, then the diameter of D_i may be supposed arbitrarily small. We may now indicate swiftly how C' may be "mapped" on a plane circle S' and its interior S: we suppose that, in some euclidean plane, S' is the set of points $x^2 + y^2 = 1$. We may map J on S' homeomorphically and further we may let an arbitrary point O of C' correspond to the origin of coördinates of the plane, the center of S'. Let us take an arbitrary $\epsilon > 0$, such moreover that O is at a distance greater than ϵ from J. Then we have seen that there exists a δ, $\epsilon > \delta > 0$, such that if ab is any spanning arc one of the domains complementary to it, in C', is of diameter less than ϵ: and this is necessarily the one which does not contain O. Of course, ab cannot contain O. There is an integer n such that if $a_1, \cdots, a_n$ are the points on J corresponding to the points (in polar coördinates) on S': $(2\pi/k, 1)$, $k = 1, 2, \cdots, n$, then every arc $a_i a_{i+1}$ of J is of diameter less than δ. Now, first, if P is the "polygon" of the preceding section, D the corresponding domain, it is clear that P may be mapped on the regular polygon in $\bar{S}$ preserving the correspondance already fixed for the vertices and this

* We are suppressing details, both at this point and in all of the sequel.

"mapping" may be extended to D and the interior of the regular polygon in S, since $\bar{D} = D + P$ is a closed 2-cell. But, second, the part of C' not mapped on $\bar{S}$ falls into a finite set of components D_i each of diameter less than ϵ. The boundaries of these domains D_i, i. e. the simple closed curves $a_i a_{i+1} a_i$ (the first being an arc of M, the second of J), are already mapped on a chord and an arc of S', respectively. We have to extend this correspondance to the domains D_i and the corresponding "area" in S. Then we have simply to repeat the entire construction of the paragraph above, for each D_i and a new and smaller ϵ', and to continue this inductively and indefinitely, always preserving correspondances already won. By a passage to the limit, we have the desired mapping of C' on $\bar{S}$: i. e. C' is a 2-cell. We remarked that if C' is a 2-cell, C with which we began our discussion, must also be a 2-cell. That should now be obvious, and the proof of our theorem concluded.

5. We permit ourselves one final remark. Entirely by the methods of this paper, and in particular of its first sections, it is possible to establish the following curious theorem:

There does not exist a continuous curve (even locally compact) which contains a pair of points, say x and y, such that every arc xy of this continuous curve separates it irreducibly between some pair of points.

PRINCETON.

ON THE CONTINUED FRACTIONS ASSOCIATED WITH, AND CORRESPONDING TO, THE INTEGRAL $\int_a^b \frac{p(y)}{x-y}\,dy$.

By J. SHOHAT (JACQUES CHOKHATE).

Introduction. Let $p(x)$ be integrable † and non-negative in the *finite* interval (a, b) reduced—without loss of generality in the discussion which follows—to $(-1, 1)$, with $\int_a^b p(x)\,dx > 0$. Consider the "corresponding" (C) and the "associated" (A) continued fractions arising in the development

$$
(1)\qquad
\begin{aligned}
\int_0^1 \frac{p(y)\,dy}{x-y} &\equiv \frac{b_1/}{/x} - \frac{b_2/}{/1} - \frac{b_3/}{/x} - \frac{b_4/}{/1} - \cdots (\equiv C)\\
&= \frac{\lambda_1/}{/x-c_1} - \frac{\lambda_2/}{/x-c_2} - \cdots (\equiv A)\\
&(b_i, \lambda_i, c_i = \text{const.};\ 0 < b_i < c_i < 1;\ \lambda_i > 0),
\end{aligned}
$$

and the denominators of the successive convergents of (A)—the orthogonal Tchebycheff polynomials

$$(2)\qquad \Phi_n(x; p) \equiv \Phi_n(x) = x^n - S_n x^{n-1} + \cdots \qquad (n = 0, 1, 2, \cdots)$$

which may be normalized:

$$(3)\qquad \phi_n(x; p) \equiv \phi_n(x) \equiv a_n(p)\Phi_n(x), \text{ with } \int_0^1 p(x)\phi_m(x)\phi_n(x)\,dx = \delta_{mn}$$
$$(m, n = 0, 1, 2, \cdots;\ a_n(p) \equiv a_n > 0).$$

The question as to the asymptotic behavior, for $n \to \infty$, of $a_n, b_n, c_n, \lambda_n, S_n, \cdots$ has been investigated by G. Szegö [1] and the writer [2]. It was found in case of $p(x)$ being an "S-function," i. e.

$$(S)\qquad \int_0^1 \frac{\log p(x)\,dx}{[x(1-x)]^{1/2}} \quad \text{exists:}$$

$$
(4)\qquad
\begin{aligned}
&a_n = 4^n A(1 + 0(1)),\quad b_n \to 1/4,\quad \lambda_n \to 1/16,\quad c_n \to 1/2,\\
&S_n = n/2 + s + 0(1) \qquad (n \to \infty),
\end{aligned}
$$

where A and s depend on the nature of $p(x)$ only.

The object of this paper is to throw some light on the behavior of $a_n, b_n, \lambda_n, \cdots$ in case the condition (S) is not satisfied, more precisely, in case $p(x)$ vanishes on a set of points $E_1 \subset (0, 1)$ of positive measure.

† Integration is taken in the sense of Lebesgue throughout this paper.

The underlying method is due, in part, to Faber [3]. In the present investigation we deal with a more general $p(x)$, also with b_n, c_n, $\lambda_n, \cdots$ (not with a_n only). It shows once more (Cf. [3, 4]) the intimate connection between the orthogonal polynomial $\Phi_n(x;p)$ and the Tchebycheff polynomial

$$\Pi_n(x) = x^n + \cdots \tag{5}$$

"the least deviating from zero" on the complementary set $E = C(E_1)$. [For brevity, $\Pi_n(x)$ is called the "T-polynomial" corresponding to the set E.] More generally, we show the close relation between $\Pi_n(x)$ and the polynomial $L_{n,k}(x) = x^n + \cdots$ minimizing $\int_0^1 p(x) \mid x^n + G_{n-1}(x)|^k\, dx$ $(k \geqq 1)$.

Notations. $G_m(x) = \sum_{i=0}^{m} g_i x^i$ — *arbitrary* polynomial of degree $\leqq m$; N, ϵ—arbitrarily large and arbitrarily small resp., but fixed, positive quantities; τ, σ—fixed positive quantities independent of x and n (N, ϵ, τ, σ are properly chosen in each case);

$$A_n(p) \equiv A_n = 1/a_n^2(p); \qquad B_n(p) \equiv B_n = 1/a_n^2(xp). \tag{6}$$

1. *Some preliminary formulae.* They will be used in the discussion which follows.

$$b_{2n+2} = A_n/B_n, \quad b_{2n+1} = B_{n-1}/A_n, \quad \lambda_{n+2} = A_n/A_{n+1} = b_{2n+2}\, b_{2n+3},$$
$$c_{n+1} = b_{2n+2} + b_{2n+1} = S_{n+1} - S_n \qquad (n \geqq 0),$$

$$A_n = 1/\prod_{i=1}^{(n+1)} \lambda_i = 1/\prod_{i=1}^{(2n+1)} b_i = \Delta_n/\Delta_{n+1} \tag{7}$$ [2]

$$(n \geqq 0;\ \Delta_n = \| \alpha_{i+j} \|_{i,j=0}^{n-1},\ n = 1, 2, \cdots;\ \Delta_0 = 1),$$

$$S_n = \sum_{i=1}^{n} c_i = \sum_{i=2}^{2n} b_i \qquad (n \geqq 1;\ \alpha_n = \int_0^1 p(x) x^n dx).$$

The following minimum property of $a_n(p)$ is of fundamental importance:

$$1/A_n(p) = \int_0^1 p(x)\Phi_n^2(x)\,dx = \min \int_0^1 p(x)\,[x^n + G_{n-1}(x)]^2\, dx, \tag{8}$$

which leads to

$$a_n(p) > 2^{2n-1}/\alpha_0^{1/2} \qquad (\alpha_0 = \int_0^1 p(x)\,dx) \quad (n \geqq 0) \tag{9}$$ [2]

$$\lambda_n(p) < 1/4; \quad A_n < B_n < A_{n+1} \tag{10}$$

—*inequalities holding true for any* $p(x)$.

2. *On the T-polynomial corresponding to a given set of points.* We make use of the following results from the Theory of Approximation [5, 6, 7].

(i) To any infinite bounded closed set of points M in the complex x-plane there corresponds one and only one T-polynomial $\Pi_n(x)$, i. e.

$$(11)\qquad E_n = \max \mid \Pi_n(x) \mid \leqq \max \mid x^n + G_{n-1}(x) \mid \quad (x \text{ in } M),$$

equality sign taking place if and only if $x^n + G_{n-1}(x) \equiv \Pi_n(x)$.

(ii) $\lim\limits_{n\to\infty} E_n^{1/n}$ exists and $= \rho(M) \equiv \rho$, with $0 \leqq \rho < \infty$.

(iii) The above limit ρ, characteristic for the given set M, is in a remarkable manner related to another characteristic constant for M—its "transfinite diameter" $d(M) \equiv d$ defined as follows [7]:

$$(12)\qquad d(M) \equiv d = \lim_{n\to\infty} d_n, \quad d_n^{n(n-1)/2} = \max \mid \prod_{i,j=1\,(i<j)}^{n} (x_i - x_j) \mid \ (n \geqq 2),$$
$$\text{with } 0 \leqq d < \infty,$$

where $x_1, x_2, \cdots, x_n$ range independently over M. In fact,

$$(13)\qquad \rho\ (\equiv \lim_{n\to\infty} \sqrt[n]{E_n}) = d\ (\equiv \lim_{n\to\infty} d_n).$$

(iv) Consider the special important case, where the complementary set $C(M)$ is a simply connected region (D) containing the point $x = \infty$. Let $z = \psi(x)$ effect the conformal representation of (D) on the outer region of the circle $\mid z \mid = 1$, so that $x = \infty \mid z = \infty$. Then, introducing the inverse function of z

$$(14)\qquad x = \phi(z) = tz + t_1/z + \cdots \qquad (t \neq 0):$$

$$(15)\qquad \rho(M) = d(\mu) = \mid t \mid.\dagger$$

3. *The relation of* $\Pi_n(x)$ *to certain extremal polynomials.* Hereafter $p(x)$ in (1) will be subject to the following conditions.

Conditions (P): (i) $p(x) = 0$ *over a set* $E_1 \subset (0,1)$ *of positive measure (necessarily* < 1, *for* $\int_0^1 p(x)dx > 0$), *such that the complementary set* $E = C(E_1)$ *consists of a finite number of intervals, each of length* $\geqq h > 0$. (ii) $p(x)$ *has in* E *a finite number of zeros* $x_1, x_2, \cdots, x_n$, *such that in sufficiently small intervals* I_ϵ *of length* 2ϵ $(x_i - \epsilon,\ x_i + \epsilon)$ ‡ $p(x) \mid x - x_i \mid^{-k_i} > A$, *with certain finite* $A (> 0)$, $k_i (> 0)$. (iii) $p(x) \geqq p_0 > 0$ *almost everywhere in* E *outside the above intervals* I_ϵ.

† Illustration: M is the interval $(-1, 1)$. Here: $x = \phi(z) = \frac{1}{2}(z + 1/z)$, $t = \frac{1}{2}$;

$$\Pi_n(x) = \frac{1}{2^{n-1}} \cos (n \arccos x), \quad \rho = \lim_{n\to\infty} \sqrt[n]{1/2^{n-1}} = \tfrac{1}{2} = t.$$

‡ With obvious modifications, if x_i is a boundary point of E.

Let

$$m_{n,k}(p) = \min \int_0^1 p(x) \mid x^n + G_{n-1}(x) \mid^k dx = \int_0^1 p(x) \mid L_{n,k}(x) \mid^k dx$$

$$(16) \quad L_{n,k}(x) = x^n + \cdots \qquad (k \geqq 1).\dagger$$

By virtue of the conditions (P), we can construct a polynomial $\Pi(x)$ of finite degree such that

$$(17) \quad \Pi(x) = \prod_{i=1}^{m} (x - x_i)^{2k'_i} (k'_i\text{—positive integer} \geqq k_i/2),$$
$$p(x)/\Pi(x) \geqq \tau > 0 \text{ in } E.$$

Hence, by the very definition of $m_{n,k}(p)$, $m_{n,k}(\Pi)$:

$$(18) \quad \begin{aligned} m_{n,k}(p) &> \tau m_{n,k}(\Pi) > \tau \cdot \min \int_E \Pi(x) \mid x^n + G_{n-1}(k) \mid^k dx \\ m_{n,k}(p) &\leqq \int_E p(x) \mid \Pi_n(x) \mid^k dx \leqq \alpha_0 E_n{}^k. \end{aligned}$$

$$(19) \quad m_{n,k}(\Pi) \leqq t\alpha_0 E_n{}^k.$$

$$(20) \quad m_{n,k}(\Pi) = t\alpha_0 \theta_n{}^n E_n{}^k \ (0 \leqq \theta_n < 1).$$

LEMMA. $\lim\limits_{n\to\infty} \theta_n = 1.$

Proof. Since $\overline{\lim\limits_{n\to\infty}}\ \theta_n \leqq 1$, it is sufficient to show that

$$(21) \quad \underline{\lim}_{n\to\infty}\ \theta_n = 1.$$

Assume the contrary holds, i. e.

$$\underline{\lim}_{n\to\infty}\ \theta_n = \alpha < 1.$$

Then, for infinitely many $n = n_\nu$ $(\nu = 1, 2, \cdots)$

$$(22) \quad \theta_n < \alpha + \epsilon < \beta < 1; \ m_{n,k}(\Pi) < t\alpha_0 \beta^n E_n{}^k.$$

Let $L^*_{n,k}(x) = x^n + \cdots$ represent the polynomial realizing the minimum $m_{n,k}(\Pi)$, with

$$(23) \quad l^*_{n,k} = \max \mid L^*_{n,k}(x) \mid = \mid L^*_{n,k}(\xi^*) \mid \qquad (x \text{ and } \xi^* \text{ in } E).$$

According to the conditions (P), ξ^* belongs to a certain interval I_{ξ^*} of length $\geqq h$, lying wholly in E, so that, by the Markoff-Bernstein Theorem,

$$(24) \quad \mid (d/dx) L^*_{n,k}(x) \mid \leqq \tau_0 n^2 l^*_{n,k} \qquad (\tau_0 = 2/h;\ x \text{ in } I_{\xi^*}).$$

We now make use of a device due to Dunham Jackson, and write, choosing $n > N$, so that $\delta_n = 1/4\tau_0 n^2 < h$:

† For the existence of one at least $L_{n,k}(x)$ cf. [8].

$$(25)\quad \begin{aligned} &| L^*_{n,k}(x) - L^*_{n,k}(\xi^*) | < \tau_0 n^2\, l^*_{n,k}\, | x - \xi^* | < \tfrac{1}{4} l^*_{n,k} \\ &x \text{ in the interval (one-sided with respect to } \xi^*) \\ &I_n\colon\ | x - \xi^* | = \delta_n = (1/4\tau_0 n^2)\,(I_n \subset I_{\xi^*}). \end{aligned}$$

$$(26)\qquad | L^*_{n,k}(x) | > \tfrac{3}{4} l^*_{n,k} \qquad (x \text{ in } I_n).$$

We notice that δ_n—*length of* I_n—*does not depend on* ξ^*. (26) leads to

$$(27)\qquad m_{n,k}(\Pi) = \int_E \Pi(x)\, | L^*_{n,k}(x) |^k\, dx > \tau (l^*_{n,k})^k \int_{I_n} \Pi(x)\, dx.$$

Introduce m non-overlapping intervals $I_1, I_2, \cdots, I_m$ contained in E, of length 2ϵ, separated by distances $\geqq \epsilon$, such that each I_i contains one only x_i—zero of $p(x)$ (and $\Pi(x)$) in E,—as its mid-, or if impossible, end-point (and boundary point of E). We further choose $n > N$ in (25) so that $\delta_n < \frac{1}{2}\epsilon$, and I_n cannot belong to *two* intervals I_i. Two cases must be considered.

I case. $\xi^* = x_j\ (1 \leqq j \leqq m)$; hence $I_n \subset I_j$. Here

$$(28)\quad \begin{aligned} &\Pi(x) = (x - x_j)^{2k'_j} \prod_{i=1\ (i \neq j)}^{m} (x - x_i)^{2k'_i} > \tau (x - x_j)^{2k'_j} \quad (x \text{ in } I_j) \\ &\int_{I_n} \Pi(x)\, dx > \tau \int_{I_n} (x - \xi^*)^{2k'_j}\, dx > \tau / n^\sigma. \end{aligned}$$

II case. $\xi^* \neq x_i\ (i = 1, 2, \cdots, m)$. (i) I_n is outside all I_i. Then,

$$| x_i - \xi^* | \geqq \epsilon/2,\quad | x - x_i | \geqq \epsilon/2 \qquad (i = 1, 2, \cdots, m;\ x \text{ in } I_n)$$

$$(29)\qquad \int_{I_n} \Pi(x)\, dx = \int_{I_n} \prod_{i=1}^{m} (x - x_i)^{2k'_i}\, dx > \tau \int_{I_n} dx > \tau / n^\sigma.$$

(ii) I_n belongs to a certain $I_j\ (1 \leqq j \leqq m)$. Here

$$| x - x_i | > \tau \qquad (1 \leqq i \leqq m;\ i \neq j;\ x \text{ in } I_n)$$

$$(30)\qquad \int_{I_n} \Pi(x)\, dx > \tau \int_{I_n} (x - x_j)^{2k'_j}\, dx.$$

If I_n lies to the right of x_j, then,

$$\begin{aligned} &\xi^* - \delta_n \geqq x_j, \qquad \int_{I_n} (x - x_j)^{2k'_j}\, dx \\ &\quad = (1/2k'_j + 1)[(\xi^* - x_j)^{2k'_j+1} - (\xi^* - \delta_n - x_j)^{2k'_j+1}] \geqq \tau \delta_n^{2k'_j} > \tau / n^\sigma, \end{aligned}$$

and the same result holds, if I_n is to the left of x_j. Finally, if x_j is inside I_n, say: $\xi^* < x_j < \xi^* + \delta_n$, then

$$\int_{I_n} (x - x_j)^{2k'_j} dx = [1/(2k'_j + 1)]\{[\delta_n - (x_j - \xi^*)]^{2k'_j+1} + (x_j - \xi^*)^{2k'_j+1}\}.$$

Whether $x_j - \xi^* \leqq \delta_n/2$ or $> \delta_n/2$, the above expression is greater than $[1/(2k'_j+1)](\delta_n/2)^{2k'_j+1} > \tau n^{-\sigma}$, and the same result holds, if $\xi^* - \delta_n < x_j < \xi^*$. Hence, in all cases

$$(31) \qquad \int_{I_n} \Pi(x)\,dx > \tau/n^\sigma,$$

so that, by (26, 22),

$$m_{n,k}(\Pi) = \int_E \Pi(x)\,|\,L^*_{n,k}(x)\,|^k\,dx > (3l^*_{n,k}/4)^k \int_{I_n} \Pi(x)\,dx > \tau(l^*_{n,k})^k/n^\sigma$$

$$(32) \qquad \alpha_0\beta^n E_n^k > \tau(l^*_{n,k})^k/n^\sigma \text{ for infinitely many } n,$$

which leads, since $0 < \beta < 1$, to the following inequality—*impossible by the very definition of* E_n:

$$E_n > l^*_{n,k} \text{ for infinitely many } n.$$

Our Lemma is thus established, and (20) leads to

$$(33) \qquad \lim_{n\to\infty} \sqrt[n]{m_{n,k}(\Pi)} = \lim_{n\to\infty} [\sqrt[n]{t\alpha_0}\,\theta_n \sqrt[n]{E_n^k}] = \rho^k.$$

Moreover, combining (33, 17, 18):

$$\sqrt[n]{\tau \cdot m_{n,k}(\Pi)} < \sqrt[n]{m_{n,k}(p)} \leqq \sqrt[n]{\alpha_0 E_n{}^k}$$

$$(34) \qquad \lim_{n\to\infty} \sqrt[n]{m_{n,k}}\,(p) = \rho^k.$$

We can go further and establish an asymptotic relation between E_n and

$$(35) \qquad l_{n,k} = \max |\,L_{n,k}(x)\,| \text{ in } E.$$

Introducing $x = \xi$ such that

$$l_{n,k} = |\,L_{n,k}(\xi)\,|$$

and the interval I'_n: $|\,x - \xi\,| = \delta_n$, with $\delta_n = 1/4\tau_0 n^2$, as in (25), we get, in view of (17):

$$(36) \qquad \int_{I'_n} p(x)\,dx > \tau \int_{I'_n} \Pi(x)\,dx > \tau/n^\sigma.$$

Hence (see (18))

$$(37) \qquad \alpha_0 E_n{}^k \geqq m_{n,k}(p) > \tau l^k{}_{n,k} n^{-\sigma}; \quad \tau E_n{}^k n^\sigma > l^k{}_{n,k} \geqq E_n{}^k$$

$$(38) \qquad \lim_{n\to\infty} \sqrt[n]{l_{n,k}} = \lim_{n\to\infty} \sqrt[n]{E_n} = \rho.$$

Our analysis thus leads to the following

THEOREM I. *Let* $p(x)$ *satisfy the conditions* (P). *Denote by* $L_{n,k}(x)$

$= x^n + \cdots$ *the polynomial*† *minimizing* $\int_0^1 p(x) \mid x^n + G_{n-1}(x) \mid^k dx$ $(k \geqq 1)$, *with* $m_{n,k}(p) = \int_0^1 p(x) \mid L_{n,k}(x) \mid^k dx$ *and* $l_{n,k} = \max \mid L_{n,k}(x) \mid$ *on the set* E *which enters into conditions* (P). *Denote further by* $\Pi_n(x) = x^n + \cdots$ *the T-polynomial corresponding to the set* E, *with* $E_n = \max \mid \Pi_n(x) \mid$ *in* E. *Then*

$$\lim_{n\to\infty} \sqrt[n]{m_{n,k}(p)} = \rho^k; \quad \lim_{n\to\infty} \sqrt[n]{l_{n,k}} = \rho = \lim_{n\to\infty} \sqrt[n]{E_n}.$$

Our analysis also yields the following theorem which will prove useful when dealing with problems similar to that under discussion.

THEOREM II. *$p(x)$ satisfying the conditions (P), if x and $x + \delta$ are any two points in E, with $\mid \delta \mid \leqq h$, then*

$$(39) \qquad \left| \int_x^{x+\delta} p(t)\,dt \right| > \tau \mid \delta \mid^\sigma \qquad (\sigma > 0),$$

where τ, σ are certain fixed positive quantities independent on x and δ. In particular, if $p(x)$ is a polynomial,

$$\left| \int_x^{x+\delta} p(t)\,dt \right| > \tau \mid \delta \mid^\sigma \qquad (\sigma > 0;\ a \leqq x,\ x + \delta \leqq b),$$

where τ, σ are certain finite numbers independent on δ or x, the latter being confined to an arbitrarily chosen, but fixed, finite interval (a, b).

The above analysis shows that *Theorem I holds for any $p(x)$ satisfying the relation (39) on the set E as given in the conditions (P).*

4. *Case $k = 2$. Application to orthogonal Tchebycheff polynomials.* In the preceding discussion take $k = 2$ and use (7, 8). We derive

THEOREM III. *If $p(x)$ satisfies the conditions (P) (more generally, if (39) is satisfied), then,*

$$(40) \qquad \sqrt[n]{1/A_n} \to \rho^2, \quad \sqrt[n]{F_n(p)} \to \rho, \quad \sqrt[n]{f_n(p)} \to 1$$
$$(\rho = \lim \sqrt[n]{E_n} = d(E);\ \tfrac{1}{4} \geqq \rho \geqq h/4)$$

$F_n(p) = \max \mid \Phi_n(x;\ p) \mid$, $f_n(p) = \max \mid \phi_n(x;\ p) \mid$ in E.

$(n \to \infty)$

$$(41) \quad \sqrt[n]{\lambda_n} \to 1, \quad \sqrt[n]{b_n} \to 1, \quad \sqrt[n]{c_n} \to 1, \quad \sqrt[n]{S_n} \to 1;\ \ \underline{\lim}\, \lambda_n \leqq \rho^2 \leqq \overline{\lim}\, \lambda_n.$$

The inequalities $\frac{1}{4} \geqq \rho \geqq h/4$ follow from our hypothesis that $E \subset (0, 1)$,

† $L_{n,k}(x)$ denotes a minimizing polynomial, if the latter is not unique.

of measure < 1, consists of intervals of length $\geqq h$.† We get the limiting relations $\sqrt[n]{c_n} \to 1$, $\sqrt[n]{S_n} \to 1$ $(n \to \infty)$, making use of:

$c_n = b_{2n-1} + b_{2n} < 1$, $\sqrt[n]{a+b} > \sqrt[n]{2}\,(\sqrt[n]{a}\cdot\sqrt[n]{b})^{1/2}$ $(a, b > 0)$, $n > S_n > c_n$ (S_n—sum of the zeros of $\phi_n(x)$—all between 0 and 1).

5. *Further discussion of the asymptotic behavior of a_n, λ_n, b_n.* Since the condition (S) is not satisfied, we know that

$$A_n \cdot 2^{-4n} \to \infty \qquad (n \to \infty) \; [1, 2]. \tag{42}$$

If $\epsilon_n, \epsilon'_n, \cdots$ denote positive quantities $\to 0$ with $1/n$, we can express the aforegoing results as follows:

$$\begin{gathered} A_n = \rho^{-2n}(1+\epsilon_n)^{-2n}, \quad B_n = \rho^{-2n}(1+\epsilon'_n)^{-2n}, \quad b_{2n+2} = \left(\frac{1+\epsilon'_n}{1+\epsilon_n}\right)^{2n}, \\ b_{2n+1} = \rho^2\left(1+\frac{\epsilon_n}{1+\epsilon'_{n-1}}\right)^{2n}(1+\epsilon^2_{n-1}), \quad \lambda_n = \rho^2\left(\frac{1+\epsilon_{n+1}}{1+\epsilon_n}\right)^{2n}(1+\epsilon_{n+1})^2. \end{gathered} \tag{43}$$

(42, 43), combined with the fact that $2\rho \leqq \frac{1}{2}$, lead to

THEOREM IV. *Assuming the existence of the limiting relations*

$$\lim_{n\to\infty}(1+\epsilon_n)^{-2n} = A^2, \qquad \lim_{n\to\infty}(1+\epsilon'_n)^{-2n} = B^2,$$

we have necessarily: $A = B = \infty$, or $A = B = 0$, save the case where the set E consists of one single interval.‡

In fact, the inequalities (10) show that if one of the quantities A, B is 0 or ∞, so is the other. Let $0 < A, \mathrm{B} < \infty$. Then,

$$\begin{gathered} A_n = \rho^{-2n}A^2(1+0(1)), \; B_n = \rho^{-2n}B^2(1+0(1)), \; \lambda_n \to \rho^2 \\ b_{2n+1} = ((B/A)\rho)^2 \equiv b', \; b_{2n} \to (A/B)^2 \equiv b'', \; c_n \to b' + b'' \\ (n \to \infty). \end{gathered} \tag{44}$$

By a theorem of Blumenthal [10], the zeros of $\phi_n(x)$ are everywhere dense (for $n \to \infty$) in $[(b' - b'')^2, (b' + b'')^2]$, while, on the other hand

† $m \subset M$ implies $d(m) \leqq d(M)$; M is a segment of length l, say, $(-l/2, l/2)$, implies $d(M) = l/4$ [9]. The latter assertion follows at once from (13), if we recall that for such M, $\Pi_n(x) = \frac{l^n}{2^{2n-1}}\cos\left(n \arccos \frac{2x}{l}\right)$, $E_n = \frac{l^n}{2^{2n-1}}$.

‡ In this case (0, 1) is reduced to a subinterval in which the condition (S) and all it implies hold true. If, for example, E is $(a, 1)$ $(0 < a < 1)$, then:

$$A_n = \frac{4^{2n-1}}{(1-a)^{2n}} \cdot \alpha[1 + 0(1)] \;\; (0 < \alpha < \infty \;\; [2]) = \rho^{-2n}A^2[1 + 0(1)],$$

with $0 < A < \infty$, since $\rho = \frac{1-a}{4}$ (See footnote †, above).

[11], if $\int_a^b p(x)dx = 0$ $(0 < \alpha < \beta < 1)$, $\phi_n(x)$ can have in (α, β) at most one zero.

Remark. In case E is the interval $(0,1)$, $\rho = \frac{1}{4}$, and (44) reduces to the asymptotic relations (4). The latter can be obtained in a very elementary way, by making simple use of the orthogonality properties in (3). Assume, for example,

$$p(0)p(1) \textit{ is finite and } \neq 0,\ p'(x) \textit{ is integrable in } (0,1).$$

(3) yields at once:

$$p(x)\phi_n^2(x)\Big|_0^1 = \int_0^1 p'(x)\phi_n^2(x)\,dx$$

$$p(x)x\phi_n^2(x)\Big|_0^1 = \int_0^1 p'(x)x\phi_n^2(x)\,dx + 2n + 1 \tag{45}$$

$$p(x)\phi_n(x)\phi_{n+1}(x)\Big|_0^1 = \int_0^1 p'(x)\phi_n(x)\phi_{n+1}(x)\,dx + (n+1)/\lambda_{n+2}^{1/2}$$

Hence *if* $p'(x)/p(x)$ *is bounded in* $(0,1)$:

$$\phi_n(c) = \sqrt{\frac{2n+1}{p(c)}} + O(1) \qquad (c = 0, 1)$$

$$\lambda_{n+2} = \frac{A_n}{A_{n+1}} \leq \frac{(n+1)^2}{4(2n+1)(2n+3) + O(1)} = \frac{1}{16} + O\left(\frac{1}{n^2}\right)$$

$$A_n = 4^{2n}A^2(1 + O(1/n)) \qquad (0 < A < \infty).$$

Introduce

$$p^*(x) \equiv p(x^2)\,|\,x\,| \text{ corresponding to } (-1, 1). \tag{46}$$

Then, as it readily follows from the orthogonality properties:

$$\begin{aligned} &a_n(p) = a_{2n}(p^*), \quad a_n(xp) = a_{2n+1}(p^*), \quad b_n(p) = \lambda_n(p^*) \\ &\text{i.e.}\ \ c_n(p) = \lambda_{2n-1}(p^*) + \lambda_{2n}(p^*), \quad \lambda_n(p) = \lambda_{2n-2}(p^*)\lambda_{2n-1}(p^*). \end{aligned} \tag{47}$$

$p^*(x)$ satisfies conditions of the type (P), with E_1, E, ρ replaced now by E^*_1, E^*, ρ^* resp. Thus, as above:

$$A^*_n \equiv A_n(p^*) = \rho^{*-2n}(1 + \epsilon^*_n)^{-2n}, \quad \lambda_n(p^*) = p^{*2}\left(\frac{1 + \epsilon^*_{n+1}}{1 + \epsilon^*_n}\right)(1 + \epsilon^*_{n+1})^2$$

$$\lambda^*_n \equiv \lambda_n(p^*) \to 1, \quad \sqrt[n]{1/A^*_n} \to p^{*2}, \quad \varliminf \lambda^*_n \leqq \rho^{*2} \leqq \varlimsup \lambda^*_n \qquad (n \to \infty,\ \epsilon^*_n \to 0). \tag{48}$$

On the other hand, by (47),

$$A_{2n}(p^*) = \rho^{-2n}(1 + \epsilon_n)^{-2n}, \quad A_{2n+1}(p^*) = \rho^{-2n}(1 + \epsilon'_n)^{-2n}, \tag{49}$$

and we thus incidentally derive the *following relation between the transfinite diameters of the two sets* $E^* = C(E^*_1)$, $E = C(E_1)$, *the correspondence of which is given by* (46):

(50) $$\rho^* = \rho^{1/2}.$$

Can we have a definite limiting value for $\lambda_n(p^*)$, *as* $n \to \infty$? We see at once, that

(51) $$\lim_{n\to\infty} \lambda_n(p^*) = \lambda \text{ implies } \lambda^* = \rho^{*2} = \rho \ \left(\text{i. e.} \left(\frac{1+\epsilon^*_{n+1}}{1+\epsilon^*_n}\right)^{2n} \to 1\right);$$

also, by (47),

(52) $$b_n(p) \to \rho, \quad \lambda_n(p) \to \rho^2, \quad c_n(p) \to 2\rho \qquad (n \to \infty),$$

which, as we have seen, is impossible. Hence, the conclusion: $\lim_{n\to\infty} \lambda_n(p^*)$ *does not exist,*† *save in the exceptional case indicated above.*

6. *Expansion of functions in series of Tchebycheff polynomials.*

THEOREM V. *If* $p(x)$ *satisfies the condition* (P), *the expression*

$$f(x) \sim \sum_{n=0}^{\infty} A_n \phi_n(x) \qquad \left(A_n = \int_0^1 p(x) f(x) \phi_n(x)\, dx\right)$$

converges absolutely and uniformly to $f(x)$ *over* E, *provided* $f(x)$ *is analytic in* $(0,1)$.

Proof. It is known from the Theory of Approximation that for such $f(x)$ we can construct a polynomial $P_n(x)$, of degree $\leqq n$, such that

$$\max |f(x) - P_n(x)| \text{ on } (0,1) = o(l^n), \qquad l < 1.$$

Hence, making use of the orthogonality properties and of Schwarz' inequality:

$$|A_{n+1}| = \left| \int_0^1 p(x) [f(x) - P_n(x)] \phi_{n+1}(x)\, dx \right| < \tau l^n$$

$$\overline{\lim_{n\to\infty}} \sqrt[n+1]{|A_{n+1}\phi_{n+1}(x)|} \leqq l < 1 \qquad (x \text{ in } E)$$

(by Theorem III), and this proves our statement.

7. *Extension to the complex domain.*‡ Let D represent a simply connected finite region in the complex x-plane, bounded by a closed curve C,

---

† If, for example, E is the interval $(0,1)$, then, of course, $\lim_{n\to\infty} \lambda_n(p)$ exists, so that $[(B/A)\rho]^2 = (A/B)^2 = \rho$, $B^2/A^2 = 1/\rho = 4$, $b_n \to 1/4$. It is of interest to compare this result with Szegö's formula [1]:

$$A = (2/\pi)^{1/2} \exp\left(-1/2\pi \int_0^1 \frac{\log p(x)}{[x(1-x)]^{1/2}}\, dx\right),$$

$$B = (2/\pi)^{1/2} \exp\left(-1/2\pi \int_0^1 \frac{\log [x p(x)]}{[x(1-x)]^{1/2}}\, dx\right).$$

We get, for example, letting $x = \sin^2\theta$: $\int_0^{\pi/2} \log \sin\theta \, d\theta = -(\pi/2) \log 2$.

‡ This section is the outcome of a stimulating conversation with Prof. J. L. Walsh.

without double points, consisting of a finite number of analytic arcs. The points of D and C form a point-set E, to which corresponds a definite T-polynomial $\Pi_n(x) = x^n + \cdots$, with all the properties given above (§ 2). Let $p(x)$ be continuous and positive on C (these restrictions could be greatly modified without impairing the results which follow). Using the previous notations, introduce the polynomial $L_{n,k}(x) = x^n + \cdots$ minimizing

$$(1/L) \int_C p(x) \mid x^n + \cdots \mid^k d\sigma \; (k > 1) \qquad (L\text{ — total circumference, } d\sigma\text{ — arc-element of } C),$$

with

$$l_{n,k} = \max \mid L_{n,k}(x) \mid \text{ on } C, \qquad m_{n,k} = (1/L) \int_C p(x) \mid L_{n,k}(x) \mid^k d\sigma.$$

Making use of the results of Fekete and Faber, as stated above (§ 2, formulae (11-15)), we readily prove

THEOREM VI. *The conclusions of Theorem I remain valid, under the conditions stated, in the complex domain*, i. e.

$$\sqrt[n]{m_{n,k}} \to \rho^k = \mid t \mid^k, \qquad \sqrt[n]{l_{n,k}} \to \rho = \mid t \mid \qquad (n \to \infty).$$

Proof. We have, as before (§ 3),

$$m_{n,k} \leq \alpha_0 E_n{}^k; \quad m_{n,k} = \alpha_0 \theta_n{}^n E_n{}^k \qquad (0 < \theta_n < 1),$$

and there remains to prove once more that $\theta_n \to 1$, as $n \to \infty$. Let

$$l_{n,k} = \mid L_{n,k}(\xi) \mid; \quad p(x) \geqq p_0 > 0 \text{ on } C.$$

The following important inequality, due to Szegö [17], is an extension to the complex domain of Markoff's theorem and will serve the same purpose:

$$(53) \qquad \mid G_n(x) \mid \leqq \mu \text{ on } C \text{ implies } \mid G'_n(x_0) \mid \leqq A\mu n^\alpha,\ 0 < \alpha \leqq 2 \ (x_0 \text{ on } C).$$

Here α *and* A *are independent on* n *and* x, *depending on the behavior of* C *in the neighborhood of* x_0; A *can be taken the same for all* x_0 *sufficiently close to a fixed point on* C.

Thus we get, integrating along C on an arc $\xi\xi_1$ of length $\sigma_{\xi\xi_1}$ sufficiently small:

$$(54) \qquad \mid L_{n,k}(x) - L_{n,k}(\xi) \mid = \mid \int_{\xi x} L'_{n,k}(x)\,dx \mid < A l_{n,k} n^2 \sigma_{\xi\xi_1} \ (x \text{ on } \xi\xi_1).$$

Since $\sigma_{\xi x}$—the length of a variable arc ξx—takes on continuously all values from 0 to L, we can choose n sufficiently large so that

$$\sigma_{\xi\xi_1} = 1/4An^2 < L,$$

with ξ_1 lying so close to ξ that (54) is applicable. Then

$$| L_{n,k}(x) | > \frac{3l_{n,k}}{4} \ (x \text{ on } \xi\xi_1); \ m_{n,k} \geqq \frac{1}{L} \int_{\xi\xi_1} p(x) \, | L_{n,k}(x) |^k \, d\sigma > \frac{\tau l^k_{n,k}}{n^2}.$$

Since this inequality is precisely of the same form as (32) derived above for the real case, the proof is achieved in the same way.

Special case: $k = 2$. Here [1] (see (7))

$$L_{n,2}(x) \equiv \phi_n(x)/a_n, \qquad l_{n,2} = 1/a_n^2 = D_{n+1}/D_n,$$

where $\phi_n(x) = a_n x^n + \cdots$ $(n \geqq 0;\ a_n > 0)$ stand for the orthogonal and normal system of Tchebycheff polynomials corresponding to the curve C and to the characteristic function $p(x)$, i. e. ($\bar{a}$ denoting the conjugate of a)

$$\frac{1}{L}\int_C p(x)\phi_m(x)\phi_n(x)d\sigma = \delta_{mn} \qquad (m, n = 0, 1, \cdots),$$

and D_n denotes the (positive) determinant

$$D_n = \begin{vmatrix} \alpha_{00} & \alpha_{10} & \cdot & \cdot & \cdot & \alpha_{n-1,0} \\ \alpha_{01} & \alpha_{11} & \cdot & \cdot & \cdot & \alpha_{n-1,1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_{0,n-1} & \cdot & \cdot & \cdot & \cdot & \alpha_{n-1,n-1} \end{vmatrix}$$

$$(n \geqq 0;\ D_0 = 1;\ \alpha_{pq} = \bar{\alpha}_{qp} = \frac{1}{L}\int_C p(x) x^p \bar{x}^q d\sigma).$$

Hence,

$$(55) \qquad \sqrt[n]{1/a_n} = \sqrt[2n]{D_{n+1}/D_n} \to | t |, \ \sqrt[n]{f_n} \to 1,$$
$$\text{as } n \to \infty \ (f_n = \max | \phi_n(x) | \text{ on } C),$$

which leads to

THEOREM VII. *The development*

$$f(x) \sim \sum_{n=0}^{\infty} A_n\phi_n(x) \qquad (A_n = \frac{1}{L}\int_C p(x)f(x)\overline{\phi_n(x)}d\sigma)$$

converges uniformly and absolutely to $f(x)$ in the closed domain D, provided $f(x)$ is therein analytic.

This is an analogon to Theorem V, and its proof is quite similar. In fact, by hypothesis, $f(x)$ is analytic in a certain domain $D_1 \supset D$. Hence [14], there exists a polynomial $P_m(x)$ of degree $m = p(n+1) - 1$ (p-fixed positive integer) such that

$$| f(x) - P_m(x) | < Hk^{n+1}, \quad k < 1 \qquad (x \text{ in } D)$$
$$(H, k \text{ independent on } x \text{ and } n),$$

and, as above,

$$| A_{n+1} | = \left| \frac{1}{L} \int_C p(x)[f(x) - P_m(x)] \overline{\phi_{m+1}(x)}\, d\sigma \right| < Hk^{n+1}(\alpha_0)^{1/2}$$

$$\overline{\lim_{m\to\infty}} \sqrt[m+1]{| A_{m+1}\phi_{m+1}(x) |} \leqq k^{1/p} < 1 \qquad (x \text{ in } D).$$

Remark. The cases $k = 2$, C is the unit-circle, or $p(x) \equiv 1$, C is an analytical curve, have been fully discussed by Szegö [15, 16].

BIBLIOGRAPHY.

(Referred to above)

[1] G. Szegö, "Ueber die Enwickelung einer analytischen Funktion," *Mathematische Annalen*, Vol. 82 (1921), pp. 188-212.

[2] Jacques Chokhate (J. Shohat), "Sur le dévelopment de l'intégrale $\int_a^b \frac{p(y)dy}{x-y}$," *Rendiconti Circolo Matematico di Palermo*, Vol. 47 (1923), pp. 25-46.

[3] G. Faber, "Ueber nach Polynomen fortschreitende Reihen," *Sitzungsber. Bayer. Akad. der Wiss., Phys.-Math. Klasse* (1922), pp. 157-178.

[4] S. Bernstein, "Sur les polynomes orthogonaux," *Journal des Mathématiques*, Vol. 9 (1930), pp. 127-177.

[5] Ch. de la Vallée-Poussin, "Sur les polynomes d'approximation à une variable complexe," *Bull. Acad. r. de Belgique*, Vol. 3 (1911), pp. 199-211.

[6] G. Faber, "Ueber Tchebycheffsche Polynome," *Crelle*, Vol. 150 (1919), pp. 79-106.

[7] M. Fekete, "Ueber die Verteilung der Wurzeln," *Mathematische Zeitschrift*, Vol. 17 (1923), pp. 228-249.

[8] J. Shohat, "On the polynomial and trigonometric approximation," *Mathematische Annalen*, Vol. 102 (1930), pp. 157-175.

[9] M. Fekete, "Ueber den transfiniten Durchmesser ebener punktmengen," *Mathematische Zeitschrift*, Vol. 32 (1930), pp. 108-114.

[10] O. Blumenthal, "Ueber die Entwickelung einer willkürlichen Funktion nach den Nennern" (*Thesis*, Göttingen, 1898).

[11] Stieltjes, "Recherches sur les fractions continues," *Oeuvres*, Vol. 2, pp. 402-566.

[12] G. Julia, "Sur les polynomes de Tchebicheff," *Comptes Rendus*, Vol. 182 (1926), pp. 1201-1202.

[13] G. Pólya, "Sur un algorithme toujours convergent," *Comptes Rendus*, Vol. 157 (1913), pp. 840-843.

[14] P. Montel, "Leçons sur les séries des polynomesa à une variable complexe" (*Borel's Monographs*), Paris, 1910, Ch. II.

[15] G. Szegö, "Beiträge zur Theorie der Toeplitzschen Formen," Teil IV, *Mathematische Zeitschrift*, Vol. 9 (1921), pp. 167-191.

[16] G. Szegö, "Ueber orthogonale Polynome," *Mathematische Zeitschrift*, Vol. 9 (1921), pp. 218-270.

[17] G. Szegö, "Ueber einen Satz von A. Markoff," *Mathematische Zeitschrift*, Vol. 23 (1925), pp. 45-61.

THE UNIVERSITY OF PENNSYLVANIA.

A SET OF TOPOLOGICAL INVARIANTS FOR GRAPHS.†

By Hassler Whitney.‡

1. *Introduction.* A linear graph, or let us say, a *topological graph,* $\mathfrak{Y}$, is a point set consisting of a finite number of points, or vertices, and a finite number of open arcs (topological images of an open segment) which do not intersect, joining pairs of these points. If we consider the vertices and arcs as abstract elements instead of as point sets, and name the two vertices which each arc joins, we obtain the corresponding *abstract graph* G. Corresponding to each abstract graph G we can form a topological graph $\mathfrak{Y}$.

Now we may take an arc of $\mathfrak{Y}$ and consider it as the sum of three point sets, a vertex (an inner point of the arc) and two arcs. If ab is the corresponding arc of G joining the vertices a and b, the above subdivision of the arc of $\mathfrak{Y}$ corresponds to replacing the arc ab by the two arcs ac and cb, c being a new vertex. The resulting graph G' represents the same point set $\mathfrak{Y}$.

If two abstract graphs G and G' can be reduced to the same abstract graph G^* by a process of subdivision, then they can be made to represent the same topological graph $\mathfrak{Y}$. Conversely, if they both represent a topological graph $\mathfrak{Y}$, then they can be reduced to the same abstract graph G^* by subdivision.§ It is therefore natural to define two such abstract graphs G and G' as being topologically equivalent or homeomorphic. Any number defined for all graphs which is unaltered when an arc is subdivided is called an *invariant under subdivision*; by the above remark, we may also call it a *topological invariant.*

A very simple topological invariant for a graph is the number of con-

† Presented to the American Mathematical Society, December 28, 1931. References:

I. "Non-separable and planar graphs," *Transactions of the American Mathematical Society*, Vol. 34 (1932), pp. 339-362.

II. "A logical expansion in mathematics," *Bulletin of the American Mathematical Society*, Vol. 38 (1932), pp. 572-579.

III. "Congruent graphs and the connectivity of graphs," *American Journal of Mathematics*, Vol. 54 (1932), pp. 150-168.

IV. "The coloring of graphs," *Annals of Mathematics*, Vol. 33 (1932), pp. 688-718.

‡ National Research Fellow.

§ See O. Veblen's "Colloquium Lectures," *Analysis Situs*, Ch. I, § 12.

nected pieces; another is the nullity † (cyclomatic number). In this paper we give a set of topological invariants which come from a set of numbers m_{ij} defined by the author.‡

2. *The invariants.* Given the table of the m_{ij} for a graph G, if we sum over the elements in each row with alternating signs, we get the m_i, the coefficients of the polynomial $M(\lambda)$ for the number of ways of coloring G in λ colors. Suppose, instead, we sum over the columns; we get a set of numbers p_i, which we shall show are topological invariants of the graph. The numbers are, if G is of rank R, nullity N,

$$p_i = \sum_j (-1)^{i+j} m_{R-j,N-i}. \tag{1}$$

The number of non-zero numbers p_i (if there are any) equals one plus the nullity N of G. For a graph with no arcs, $p_0 = 1$, and $p_i = 0$, $i \neq 0$.

Suppose G is planar, and has a dual G'.§ From the definition of dual graphs it follows immediately that if m'_{ij} are the numbers for G', then

$$m'_{ij} = m_{R-j,N-i}. \tag{2}$$

Hence

$$m'_i = \sum_j (-1)^{i+j} m'_{ij} = \sum_j (-1)^{i+j} m_{R-j,N-i} = p_i, \tag{3}$$

that is, *if G has a dual G', then the numbers p_i are the coefficients m'_i of $M'(\lambda)$.*

From this it follows immediately that the p_i are topological invariants for G, provided G is planar. For take any arc ab of G and replace it by the two arcs $ac + cb$, c being a new vertex, forming the graph G^*. If G' is a dual of G and $a'b'$ is the arc corresponding to ab, then we can form a dual G'' of G^* by adding to G' another arc $a'b'$, as is easily proved. Now any coloring of G' is a coloring of G'' and conversely, and hence $m'_i = m_i''$. As $p_i = m'_i$ and $p^*_i = m_i''$, it follows that $p_i = p^*_i$.

We give now a direct and general proof that $p_i = p^*_i$. Divide the subgraphs of nullity $N - i = N^* - i$ of G^* into two groups: (1) those containing the arc ac, and (2) those not containing ac. Consider first the subgraphs in (2); we pair them off, letting correspond to each subgraph H_1 not containing the arc cb, the subgraph $H_2 = H_1 + cb$. Say H_1 is of rank

† See I, § 2. If a graph G containing E arcs and V vertices is in P connected pieces, then its rank R and nullity N are defined by the equations $R = V - P$, $N = E - R = E - V + P$.

‡ See II, § 6. A subgraph H of G is determined by naming a subset of the arcs of G. m_{ij} is the number of subgraphs of G of rank i, nullity j.

The author was mistaken in supposing that the m_{ij} are the same as Birkhoff's (i, k).

§ See I, § 8.

$R^* - j$; then H_2 is of rank $R^* - j + 1$. H_1 contributes to $m^*_{R^*-j, N^*-i}$, and thus to p^*_i with the sign $(-1)^{i+j}$; H_2 contributes to $m^*_{R^*-j+1, N^*-i}$, and thus to p^*_i with the opposite sign. The two contributions cancel; thus the subgraphs in (2) together contribute nothing. We have left only the subgraphs of the first group to consider.

To each subgraph H^* of the first group we let correspond that subgraph H of G formed by dropping out the arc ac and letting the vertices a and c coalesce. (Thus the arc cb, if present, is replaced by the arc ab.) This is a $1-1$ correspondence between all the subgraphs of G of nullity $N - i$ and the subgraphs of G^* in the first group. If H^* is of rank $R^* - j$, then H is of rank $R^* - j - 1 = R - j$. Hence the contribution of H^* to p^*_i is the same as the contribution of H to p_i. It follows that $p^*_i = p_i$, as required.

3. *Broken cut sets of arcs.* We give here an interpretation of the p_i dual to the interpretation of the m_i in terms of broken circuits (see II, § 7). Suppose that dropping out a set of arcs $\alpha, \beta, \cdots, \delta$ from a graph G increases the number of connected pieces in G, while dropping out no proper subset of them does. We then say these arcs form a *cut set of arcs.* List the arcs of G in a definite order. From each cut set of arcs we drop out the last arc, forming the corresponding *broken cut set of arcs.* If G contains a cut arc, then we can consider the null subgraph of G as the corresponding broken cut set.

If G is planar and G' is a dual of G, then cut sets of arcs in G correspond to circuits in G' and broken cut sets, to broken circuits (see I, Theorem 9). As an example, let G^* be the graph with vertices a^*, b^*, c^*, and arcs $\alpha^*(a^*b^*)$, $\beta^*(a^*b^*)$, $\gamma^*(b^*c^*)$, $\delta^*(a^*c^*)$, $\epsilon^*(a^*c^*)$; then G^* has as dual the graph G given in the beginning of II, § 7. The cut sets of G^* are α^*, β^*, γ^*, and γ^*, δ^*, ϵ^*, and α^*, β^*, δ^*, ϵ^*, and the broken cut sets are α^*, β^*, and γ^*, δ^*, and α^*, β^*, δ^*. For G^*, $p^*_0 = 1$, $p^*_1 = -5$, $p^*_2 = 8$, $p^*_3 = -4$.

We now prove that $(-1)^i p_i$ *is the number of subgraphs of i arcs of G which do not contain all the arcs of any broken cut set of arcs.*

The proof follows the proof of the corresponding theorem in II. Arrange the broken cut sets $P_1, P_2, \cdots, P_\sigma$ in order so that, for any k, naming the arcs of G one by one in the given order, all the arcs of $P_1, \cdots, P_{k-1}$ have been named by the time that all the arcs of P_k have been. Arrange the subgraphs of G into sets $S_1, S_2, \cdots, S_\sigma, S_{\sigma+1}$ (some of which may be empty), putting into S_1 all those subgraphs containing no arcs of P_1, into S_k, $1 < k \leqq \sigma$, all those containing at least one arc from each of the broken cut sets $P_1, \cdots, P_{k-1}$, but containing no arc of P_k, and into $S_{\sigma+1}$, all remaining subgraphs.

Consider the subgraphs in any S_k, $1 \leqq k \leqq \sigma$. Let ab be the arc of the cut set which is dropped out in forming the broken cut set P_k. To each subgraph H_1 of S_k containing ab corresponds a subgraph H_2 of S_k not containing ab, and conversely, as ab is in none of the broken cut sets $P_1, \cdots, P_k$. Say H_1 is of rank $R-j$, nullity $N-i$. As the arcs of P_k together with the arc ab form a cut set of arcs and H_1 contains no arcs of P_k, dropping out ab from H_1 disconnects a and b, and thus H_2 is of rank $R-j-1$, nullity $N-i$. H_1 contributes to $m_{R-j,N-i}$, and thus to p_i with the sign $(-1)^{i+j}$, and H_2 contributes to p_i with the opposite sign; the contributions of H_1 and H_2 to p_i thus cancel. The subgraphs of $S_1, \cdots, S_\sigma$ contribute therefore nothing.

We have left the subgraphs in $S_{\sigma+1}$, that is, the subgraphs which contain at least one arc from each broken cut set $P_1, \cdots, P_\sigma$. Consider any such subgraph of nullity $N-i$; it contains an arc of each cut set, and it has therefore the same rank R as G;† hence it contains $R+(N-i)=E-i$ arcs, if G contains E arcs. Say there are l_i such subgraphs in $S_{\sigma+1}$; they contribute an amount $(-1)^i l_i$ to p_i. Now the subgraphs of $S_{\sigma+1}$ are exactly the complements of the subgraphs of i arcs of G which do not contain all the arcs of any broken cut set, and the theorem is proved.

From this interpretation of the numbers p_i, it is again easily seen that they are topological invariants.

Note that *for any graph containing a cut arc, and only for these, every* $p_i=0$. This corresponds to the fact that for any graph containing a 1-circuit, every $m_i=0$.

4. *Separable graphs.* Suppose that G is the union of two graphs G' and G'' which have at most a single vertex in common. Then

$$p_i = \sum_k p'_k p''_{i-k}.\ddagger \tag{4}$$

As a result, if $G_1, G_2, \cdots, G_n$ are the components of G and we know all the $p_i(G_s)$, we can calculate the $p_i(G)$. Components which are isolated vertices may of course be forgotten altogether.

To prove (4), arrange the arcs of G in a fixed order. Now the arcs of any cut set in G, hence also of any broken cut set in G, lie wholly in G' or G''. For suppose there were a cut set containing an arc ab in G' and an arc cd in G''. If the arcs of the cut set are dropped out of G, a and b are dis-

† We can reduce the rank of a graph only by dropping out all the arcs of some cut set.

‡ The same formula holds for the m_i.

connected. If we put back the arc cd, there is then a chain joining a and b. But any chain from a to b must lie wholly in G', and thus does not contain cd, a contradiction, proving the statement. Let $P_1, \cdots, P_\sigma$ be the broken cut sets in G', and $P_{\sigma+1}, \cdots, P_\tau$, those in G''; then $P_1, \cdots, P_\tau$ are those in G.

Let H' be any subgraph of G' of k arcs not containing any broken cut set, and let H'' be a similar subgraph of G'' of $i - k$ arcs. Then $H = H' \dot{+} H''$ is a subgraph of G of i arcs not containing any broken cut set. Conversely, any such subgraph H consists of two such subgraphs H' and H''. The number of such subgraphs H equals the number of such pairs $H', H'', = \sum_k (-1)^k p'_k \cdot (-1)^{i-k} p''_{i-k}$, as required.

5. *Completeness of the invariants.* It is easily seen that if two graphs G and G' are 2-homeomorphic (see the following paper), then $p_i = p'_i$. The question arises, if $p_i = p'_i$, then are G and G' 2-homeomorphic? (If they are triply connected, they would then be isomorphic.) This is not true, as is shown by the two following graphs.†

G: ab, bc, cd, dd_1, d_1e, ee_1, e_1f, fa, ag, bg, cg, dg, eg, fg, ac, d_1e_1;
G': ab, bc, cc_1, c_1d, dd_1, d_1e, ef, fa, ag, bg, cg, dg, eg, fg, ac, c_1d_1.

G and G' are evidently not 2-homeomorphic. That $p_i = p'_i$ may be seen as follows. First, each graph is a dual of itself; hence $p_i = m_i$, $p'_i = m'_i$. Form G^* from G by dropping out the vertex b and the arcs on it, and form G'' from G' by dropping out the same vertex and arcs. Then $M(\lambda) = (\lambda - 3)M^*(\lambda)$, $M'(\lambda) = (\lambda - 3)M''(\lambda)$. But G^* and G'' are isomorphic, hence $M^*(\lambda) = M''(\lambda)$, therefore $M(\lambda) = M'(\lambda)$, therefore $m_i = m'_i$, and thus $p_i = p'_i$.

The following question is as yet unanswered. If two graphs have the same m_{ij}, are they 2-isomorphic?‡ In the above example, $m_{32} = 10$ and $m'_{32} = 9$.

PRINCETON UNIVERSITY.

† This was discovered by R. M. Foster.

‡ The hypothesis that $m_{ij} = m'_{ij}$ is a greatly weakened form of the hypothesis of the theorem in the paper "2-isomorphic graphs," *American Journal of Mathematics*, Vol. 55 (1933), pp. 245-254.

ON THE CLASSIFICATION OF GRAPHS.†

By Hassler Whitney.‡

1. *Introduction.* R. M. Foster § has given an enumeration of graphs, for use in electrical theory. He uses two distinct methods, classifying the graphs according to their nullity, and according to their rank. In either case, only a certain class of graphs is listed; the remaining graphs are easily constructed from these. In the present paper we give theorems sufficient to put the first method of classification on a firm foundation.

In this method (see §§ 8 and 9), only the elementary graphs (see § 4), or graphs whose connected pieces are elementary, are listed. These graphs are most easily formed from the basic graphs (see § 5), and these, from the basic graphs of nullity one less. This manner of constructing the graphs, and in particular, the important notion of basic graphs, is due to Foster. The definition of elementary graphs and the proofs are, in general, due to the author. We assume here a knowledge of the first half of the paper I.¶

2. Following the terminology of electrical theory, we shall say that two arcs ab, bc, are *in series*, if the vertex b is on no other arc. The vertices a and c need not be distinct, but they must be distinct from b. Two arcs ab, ab, joining the same two distinct vertices, we shall say are *in parallel.*

We shall consider operations on graphs of the following types.

(1a) Replace an arc ab by two arcs ac and cb in series (c being a new vertex).

(1b) Replace two arcs ac and cb in series by a single arc ab, dropping out the vertex c.

In these operations, a and b need not be distinct.

(2) Break the graph at a single vertex into two connected pieces, or join two connected pieces at a single vertex.

† Presented to the American Mathematical Society, December 28, 1931, under the title "Basic graphs."

‡ National Research Fellow.

§ Ronald M. Foster, "Geometrical circuits of electrical networks," *Bell Telephone System Technical Publications*, Monograph B-653; also in the *Transactions of the American Institute of Electrical Engineers*, Vol. 51 (1932), pp. 309-317. See in this connection a paper by the author, "2-isomorphic graphs," *American Journal of Mathematics*, Vol. 55 (1933), pp. 245-254.

¶ For references, see the preceding paper.

By operations of this sort we can make one graph isomorphic with another if its components are respectively isomorphic with the components of the other.

(3) Suppose $G = H_1 \dotplus H_2$,† where H_1 and H_2 have the vertices a and b and no others in common, and a and b are connected in both H_1 and H_2. If $ac_1, ac_2, \cdots, ac_m$, and $bd_1, bd_2, \cdots, bd_n$ are the arcs of H_1 on a and b respectively (there is at least one arc in each set), replace these by the arcs $bc_1, bc_2, \cdots, bc_m, ad_1, ad_2, \cdots, ad_n$. We shall say simply, turn H_1 around at the vertices a and b.

Later on we shall also have to consider operations of the following type.

(4) If ab is an arc of G, drop out this arc and let the vertices a and b coalesce.

We note that none of these operations alter the nullity of the graph.

3. We define now certain relations between two graphs. If one graph is formed from another by employing at most operations designated in the first column, then we say the graphs are related as shown in the second column:

no operations	isomorphic ‡
(2)	1-isomorphic §
(2) and (3)	2-isomorphic ¶
(1a) and (1b)	homeomorphic
(1a), (1b) and (2)	1-homeomorphic
(1a), (1b), (2) and (3)	2-homeomorphic.

These relations are all reflexive, symmetric and transitive.

If two topological graphs † are homeomorphic in the topological sense, then the corresponding abstract graphs are homeomorphic in the above sense, and conversely.‖

THEOREM 1. *Any graph 2-homeomorphic with a non-separable graph of nullity* > 0 *is non-separable* (*and of nullity* > 0).

Let G be a non-separable graph of nullity > 0. Suppose first G' is

† $H_1 \dotplus H_2$ is the graph containing the arcs and vertices of both H_1 and H_2. If H_1 and H_2 have no common vertices, then $H_1 \dotplus H_2 = H_1 + H_2$. $H_1 \cdot H_2$ is that graph whose arcs and vertices are in both H_1 and H_2.

‡ See I, § 7; we formerly used the term "congruent." The operation of changing names of vertices and arcs we shall consider as trivial, and shall allow it at any time without mention.

§ The term equivalent was used in I.

¶ Equivalent in the sense of R. M. Foster.

‖ See the preceding paper.

formed from G by an operation of type (1a). If G' is separable, then $G' = I'_1 + I'_2$, $I'_1 \cdot I'_2 =$ a single vertex a, and I'_1 and I'_2 each contain an arc. G is formed from G' by replacing two arcs in series $bd + dc$ by the single arc bc. Then bd and dc both lie in I'_1 or in I'_2. For if not, then $d = a$, and thus any chain from b to c in G' passes through d. Hence b and c are joined in G only through the arc bc, and G is not cyclicly connected. But as G is of nullity > 0 and b and d are distinct, G contains at least two arcs, contradicting I, Theorem 7. Say bd and dc are both in I'_1. Replacing these by the arc bc, I'_1 goes into a graph I_1; put $I_2 = I'_2$. $G = I_1 \dot{+} I_2$ is seen to be separable, a contradiction.

The case that G' is formed from G by an operation of type (1b) is similar. Operations of type (3) obviously leave a non-separable graph non-separable. As no operations of type (2) are possible in a non-separable graph, the theorem is proved.

4. *Definitions.* A graph is called *elementary* if it is non-separable and is not 2-homeomorphic with any graph with fewer arcs. Suppose an arc α in a graph G, if dropped out, disconnects G. We then call α a *cut arc* of G. If the two arcs α and β disconnect G if dropped out, while neither is a cut arc, then we say they form a *cut pair of arcs* of G.

THEOREM 2. *If G is 2-homeomorphic with the elementary graph G', then G can be formed from G' by operations of types* (1a) *and* (3) *alone.*

If not, then form G from G', using the fewest possible number of operations of type (1b), and say $G' = G_0, G_1, G_2, \cdots, G_n = G$, are the successive graphs formed. (By the last theorem, operations of type (2) cannot occur.) We suppose G' is of nullity $N > 1$; the theorem is evident otherwise.

Say the first time an operation of type (1b) is employed is in forming G_i from G_{i-1}; the arcs in series $ac + cb$ are replaced by the arc ab. If ab is dropped out of G_i, a graph G^* is formed; let $G_i(1), \cdots, G_i(m-1)$ be its components of nullity > 0 (of which there is at least one), and $H_i(1), \cdots, H_i(p_i - 1)$, its components consisting of a single arc, if there are any (see I, Theorem 8). If we put $H_i(p_i) = ab$, then the graphs $G_i(1), \cdots, G_i(m-1)$, $H_i(1), \cdots, H_i(p_i)$ form a circuit of graphs, as is seen from the first part of the proof of I, Theorem 18.

We shall now show that for each number k, $0 \leqq k \leqq n$, we can put $G_k = G_k(1) \dot{+} \cdots \dot{+} G_k(m)$, and if $H_k(1), \cdots, H_k(p_k)$ are the arcs of $G_k(m)$, then

(a) Each graph $G_k(1), \cdots, G_k(m-1)$ is non-separable and of nullity > 0,

(b) The graphs $G_k(1),\cdots,G_k(m-1)$, $H_k(1),\cdots,H_k(p_k)$ form a circuit of graphs, and

(c) When G_k is formed from $G_{k\pm1}$, each graph $G_{k\pm1}(s)$ goes thereby into $G_k(s)$, $(s=1,2,\cdots,m)$.

This is true when G_i is formed from G_{i-1} and conversely; we shall show that it is true for G_k when it is formed from $G_{k\pm1}$. If G_k is formed from $G_{k\pm1}$ by an operation of type (1a), this is obvious. An arc of some graph $G_{k\pm1}(s)$ is replaced by a pair of arcs in series; $G_{k\pm1}(s)$ goes thus into $G_k(s)$, which is non-separable if $s<m$ (Theorem 1), and is a set of arcs if $s=m$. (b) obviously holds. Suppose an operation of type (1b) was employed; then two arcs α and β in series are replaced by a single arc γ. As each vertex of $G_{k\pm1}(j)$ is on at least two arcs of $G_{k\pm1}(j)$, $(j=1,\cdots,m-1)$ (I, Theorem 8), α and β lie in the same graph $G_{k\pm1}(s)$; we let the rest of $G_{k\pm1}(s)$ together with γ form $G_k(s)$. The other properties above are easily verified.

Suppose now G_k was formed from $G_{k\pm1}$ by an operation of type (3). Then $G_{k\pm1}=I_1\dotplus I_2$, and when I_1 is turned around at the vertices a and b, G_k is formed. If either I_1 or I_2 is contained wholly in one of the graphs $G_{k\pm1}(1),\cdots,G_{k\pm1}(m-1)$, the properties are quickly verified. Suppose not; then each graph $G_{k\pm1}(1),\cdots,G_{k\pm1}(m-1)$ is contained wholly in one of the graphs I_1, I_2. For otherwise some graph $G_{k\pm1}(s)$, $s<m$, contains arcs of both I_1 and I_2, and I_1 and I_2 each contain arcs in graphs $G_{k\pm1}(j)$, $(j\neq s)$. Following around the circuit of graphs from an arc of I_1 to an arc of I_2, keeping away from $G_{k\pm1}(s)$, a vertex, say b, is found common to I_1 and I_2 and not lying in $G_{k\pm1}(s)$. Thus I_1 and I_2 have but a single vertex a in common in $G_{k\pm1}(s)$, and this graph is separable, contrary to hypothesis. Now as each graph $G_{k\pm1}(1),\cdots,G_{k\pm1}(m-1)$, $H_{k\pm1}(1),\cdots,H_{k\pm1}(p_{k\pm1})$ lies wholly in I_1 or in I_2, the effect of the operation is merely to alter the arrangement of these graphs in the circuit of graphs, and the properties are again verified.

We can thus divide the operations forming G from G' into two groups: those altering one of the graphs $G_k(1),\cdots,G_k(m-1)$, and those changing the number of arcs in $G_k(m)$ or altering the arrangement of the graphs in the circuit of graphs. We can evidently form G from G' by first performing all the operations in the first group, and then performing all those in the second group. An operation of type (1b) occurs in the second group, in forming G_i from G_{i-1}. Now $G'=G_0$ is elementary, and thus contains no cut pair of arcs (see Theorem 6); hence $G_0(m)$ contains but a single arc: $p_0=1$. We can replace the operations in the second group by the following: replace $H_0(1)$ by p_n arcs in series; then, by operations of type (3), arrange the

graphs (including these arcs) properly in the circuit of graphs. We have thus formed G using fewer operations of type (1b), a contradiction, proving the theorem.

We can strengthen this theorem in the following one.

THEOREM 3. *Under the same conditions as in the last theorem, G can be formed from G' by employing first operations of type* (1a) *alone, then operations of type* (3) *alone.*

Take each arc of G which will be replaced by other arcs, and replace it at once by as many arcs in series as it will turn into. We then perform the operations of type (3), being careful merely to break the graph at the proper point each time.

An immediate consequence of Theorem 2 is

THEOREM 4. *Any two 2-homeomorphic elementary graphs are 2-isomorphic.*

The following theorem will be useful in later work.

THEOREM 5. *If a non-separable graph has a cut pair of arcs, and the arcs are not in series, then the four end vertices of these arcs are all distinct.*

For if ab and ac were a cut pair of arcs and there were an arc ad in the graph, a would be a cut vertex.

THEOREM 6. *A necessary and sufficient condition that a non-separable graph G of nullity > 0 be elementary is that it contain no cut pair of arcs.*

We prove first the necessity of the condition. Assuming that G is elementary, we shall show that it has no cut pair of arcs. G contains no two arcs in series, as otherwise, replacing them by a single arc gives a 2-homeomorphic graph with fewer arcs, a contradiction. Suppose G contained a cut pair of arcs ab, cd, not in series. Then these four vertices are distinct, by the last theorem. Dropping out these two arcs leaves two connected graphs H'_1 and H'_2, one containing the vertices a and c say, and the other, the vertices b and d. Put $H_1 = H'_1 \dot{+} ab$, $H_2 = H'_2 \dot{+} cd$. Turning H_2 around at the vertices b and c gives a graph G' 2-homeomorphic with G, with two arcs ab and bd in series; again, we find a 2-homeomorphic graph with fewer arcs.

To prove the sufficiency of the condition, suppose G is not elementary; then it is 2-homeomorphic with an elementary graph G'. By Theorem 2, G can be formed from G' by operations of types (1a) and (3) alone. Say the last operation of type (1a) was to replace an arc γ by two arcs in series

α and β; α and β are a cut pair of arcs in the resulting graph. But operations of type (3) leave these arcs a cut pair, and thus G contains a cut pair of arcs.

5. *Definitions.* A graph is called *cubic* if each vertex is on exactly three arcs. Any cubic elementary graph, also a 1-circuit, is called a *basic graph.* The basic graphs of nullities one, two, and three are: aa; ab, ab, ab; ab, ac, ad, bc, bd, cd. There are two, four, and fourteen basic graphs of nullities four, five, and six respectively (see Foster's paper).

THEOREM 7. *A basic graph G of nullity > 2 contains no (1- or) 2-circuits.*

For suppose G had a 2-circuit ab, ab. There is only one other arc α on a, and one other arc β on b, and neither of these is an arc ab. α and β are a cut pair of arcs, contradicting Theorem 6.

THEOREM 8. *A basic graph G of nullity > 2 is triply connected.*†

Obviously G contains at least four vertices. Suppose G could be disconnected into the two parts H_1, H_2, by dropping out the two vertices a, b. If, first, there is an arc ab in G, then there is but a single arc joining a to H_1, and a single arc joining b to H_1. These arcs form a cut pair in G, contradicting Theorem 6. If there is no arc ab, one of the graphs H_1, H_2 is joined to a by but a single arc, and one is joined to b by but a single arc. These arcs form a cut pair in G, again a contradiction.

THEOREM 9. *Any two 2-homeomorphic basic graphs G and G' are isomorphic.*

By Theorem 4, G and G' are 2-isomorphic. If they are of nullity 1 or 2, the theorem is true; we assume they are of nullity > 2, in which case they are triply connected. By III, Theorem 4, it is seen that the only operation of type (3) possible is the trivial one of turning around a single arc, which does not alter the graph. Thus G and G' are isomorphic.

6. THEOREM 10. *Let G be an elementary graph, and let the non-separable graph G' be formed from G by an operation of type* (4). *Then G' is elementary.*

For a cut pair of arcs of G' would evidently be a cut pair of arcs of G.

THEOREM 11. *Any elementary non-basic graph G which is not a single*

† See III, p. 158. A graph is triply connected if it contains at least four vertices, and is not disconnected by the omission of any one or two vertices.

arc can be formed from a basic graph G' of the same nullity by operations of type (4).

We shall show how an elementary graph G_1 can be formed from G by the inverse of an operation of type (4). Similarly an elementary graph G_2 can be formed from G_1, etc. Obviously we arrive at a basic graph G' after a finite number of steps; the inverse of these operations carries G' into G.

As G is not basic, there is a vertex a on at least four arcs $aa_1, aa_2, \cdots, aa_m$ $(m \geqq 4)$. Take a new vertex b, replace the arcs aa_1 and aa_2 by the arcs ba_1 and ba_2, and add the arc ab, giving a graph G_1; G is formed from G_1 by an operation of type (4).

G_1 is easily seen to be non-separable. If it is not elementary, it has a cut pair of arcs ab, cd (one of these must obviously be ab, as G has no cut pair of arcs); the four vertices are distinct (Theorem 5). Dropping out these arcs gives two connected graphs H_1 and H_2 containing say b and d, a and c, respectively. b is not a cut vertex of H_1, as otherwise a would be a cut vertex of G. Hence there is a chain C_1 joining a_1 and a_2 in $H_1 - b$. Similarly there is a chain C_2 joining a_3 and a_4 in $H_2 - a$.

Form G'_1 from G by adding the new vertex b', replacing the arcs aa_1 and aa_3 by the arcs $b'a_1$ and $b'a_3$, and adding the arc ab'. G'_1 is elementary. For suppose it had a cut pair of arcs ab', ef; then these four vertices are distinct, and every chain joining a to b' in $G'_1 - ab'$ must contain ef. But one of the chains C_1, C_2, say C_1, does not contain ef; thus $aa_2 + C_1 + b'a_1$ is a chain joining a and b' in $G'_1 - ab' - ef$, a contradiction, proving the theorem.

7. THEOREM 12. *If an arc ab is removed from a basic graph G, the resulting graph G^* is non-separable.*

G^* is surely connected. If it is separable, it has a cut vertex x. Let xy_1, xy_2, (xy_3) be the two, or three, arcs of G^* on x; then one of the vertices y_1, y_2, (y_3), say y_1, is joined to none of the others by a chain in $G^* - x$. Hence a chain from y_1 to x must contain the arc xy_1, that is, xy_1 is a cut arc of G^*. Thus ab, xy_1 are a cut pair of arcs of G, contradicting Theorem 6.

THEOREM 13. *Any basic graph G of nullity > 2 can be formed from a basic graph G_1 of nullity one less by replacing two arcs of G_1 by two pairs of arcs in series, and joining the two new vertices by a new arc.*

Given the basic graph G, we shall find such a graph G_1. Remove an arc ab from G, and replace the two pairs of arcs in series that are now present by single arcs α, β. (It is easily seen that the two pairs of arcs consist of four distinct arcs.)

If the resulting non-separable † graph G_1 is not basic, it has a cut pair of arcs cd, ef; the four vertices are distinct, by Theorem 5. These two arcs were present in G, i. e. neither is α or β. For otherwise, suppose for instance $cd = \alpha$ was formed from the two arcs $ca + ad$, while ef is not β. As $G_1 - cd - ef$ is in exactly two connected pieces, the end vertices of β are connected to either c or d in this graph, say to c. Then replacing β by the two original arcs touching b and adding the vertex a and the arcs $ab + ac$ leaves the graph unconnected, and forms the graph $G - ad - ef$; thus ad, ef are a cut pair of arcs in G, a contradiction. If $cd = \alpha$ and $ef = \beta$, then either ac, be, or ac, bf are a cut pair of arcs in G. This proves the statement. Moreover, the vertices a, b, c, d, e, f are all distinct, as a and b are not in G_1 and are thus distinct from c, d, e, f.

$G - ab - cd - ef$ is in two connected pieces H_1 and H_2; say H_1 contains a, c and e, and H_2, b, d and f. Neither graph has a cut arc. For suppose H_1 say had a cut arc xy. One of the vertices a, c, e, say a, is connected to neither of the others in $H_1 - xy$, as otherwise they would all be connected, and xy would be a cut arc of G. Then a is not connected to c in $G - xy - ab$, and xy, ab are a cut pair of arcs in G, a contradiction.

The next step is to show that if $a'b'$, $c'd'$, $e'f'$ are any cut triple of arcs in G, then either H_1 or H_2 contains none of these arcs. Otherwise, one of these graphs, say H_1, would contain exactly one of the arcs, say $a'b'$. Dropping out the three arcs disconnects a' and b' in G. But as H_1 has no cut arc, a' and b' are connected in $H_1 - a'b'$, and thus in $G - a'b' - c'd' - e'f'$, a contradiction.

We can now prove the theorem easily. Having assumed that $G - ab$ did not give a basic graph, we saw it contained a cut pair of arcs cd, ef, and contained graphs H_1 and H_2 as above. Let $a'b'$ be an arc of H_1. If $G - a'b'$ does not give a basic graph when the pairs of arcs in series are replaced by single arcs, it has a cut pair of arcs $c'd'$, $e'f'$, neither arc of which lies in H_2. Hence, of the two corresponding graphs H'_1, H'_2, one of them, say H'_1, is contained in $H_1 - a'b'$. Let $a''b''$ be an arc of H'_1; if $G - a''b''$ does not give a basic graph, we find a cut pair of arcs and graphs H_1'', H_2'', with H_1'' contained in $H'_1 - a''b''$, etc. As each graph $H_1^{(i)}$ contains arcs, this process must come to an end, and we find finally a graph $G - a^{(i)}b^{(i)}$ which gives a basic graph.

8. *The construction of graphs.* We can now give a standard method for

---

† See Theorems 1 and 12.

the construction of graphs, forming first the basic graphs, then the elementary graphs, then any graphs.

(1) The basic graphs of nullities 1 and 2 are known. We form successively the basic graphs of nullities $3, 4, \cdots$, as in Theorem 13.

(2) From the basic graphs of a certain nullity, we form all elementary graphs of the same nullity as in Theorem 11. Any non-separable graph we can form thus is elementary (Theorem 10). A given elementary graph may, however, be derived from different basic graphs. If we wish, we can forget all graphs formed which have 2-circuits (arcs in parallel).

(3) Taking all elementary graphs of a given nullity, we form all non-separable graphs of the same nullity by operations first of type (1a) and then of type (3) (Theorem 3). If we left out elementary graphs with 2-circuits and wish to include graphs with 2-circuits now, we must add arcs in parallel with various arcs of non-separable graphs of lesser nullity. We form finally any graph by taking non-separable graphs and letting vertices coalesce in such a manner that the graphs are the components of the final graph.

9. *The classification of graphs by nullity.* To list all graphs, even all non-separable graphs with no two arcs in series, of nullities say one to five, would be a tremendous task. It is thus natural to list only graphs of some class from which all non-separable graphs can be derived without too much difficulty. The elementary graphs form such a class. Moreover, from any group of 2-homeomorphic and thus 2-isomorphic elementary graphs, we need list but one.

As the larger part of the elementary graphs have 2-circuits,† a great saving could be effected by listing none of these graphs. But then, to form all non-separable graphs from these, we would have an added operation to perform, which would increase the difficulty greatly.

We note that an elementary graph with two arcs in parallel may not remain elementary if we drop out one of these arcs; hence we cannot form all elementary graphs from elementary graphs without arcs in parallel by merely replacing single arcs by arcs in parallel.

PRINCETON UNIVERSITY.

† See the figures in R. M. Foster's paper.

2-ISOMORPHIC GRAPHS.*

By Hassler Whitney.†

1. In the preceding paper we said that two graphs G and G' are 2-isomorphic if one can be transformed into the other by operations of the following two types: (2) The arrangement of the components in the graph is altered. (3) If $G = H_1 \dotplus H_2$, where H_1 and H_2 have just the vertices a and b in common and these vertices are connected in both H_1 and H_2, then H_1 is turned around at these vertices.

If G and G' are 2-isomorphic, then any circuit in one graph corresponds to a circuit in the other; for an operation of either type transforms any circuit into a circuit. It was shown in III, Theorem 2,‡ that if there is a $1-1$ correspondence between the arcs of two triply connected graphs so that circuits correspond to circuits,§ then the two graphs are isomorphic (we formerly used the term "congruent"). The question arises, what can be said about any two graphs in which circuits correspond to circuits? The answer is given in the following theorem. The phrase "strictly isomorphic (2-isomorphic)" means: "isomorphic (2-isomorphic), preserving the correspondence between the arcs of the graphs."

Theorem. *If there is a $1-1$ correspondence between the arcs of the two graphs G and G' so that circuits correspond to circuits, then the graphs are strictly 2-isomorphic.*

In this theorem we can replace the word "circuits" by the words "subgraphs of nullity 0" or "subgraphs of nullity 1" or "cuts sets of arcs" (see the paper on topological invariants). For the first statements see the proof of III, Theorem 3; for the last, see a paper "Planar graphs."

As an example, the graphs G' and G'' of I, p. 353, are strictly 2-isomorphic.

---

* Presented to the American Mathematical Society, August 30, 1932.

† National Research Fellow.

‡ See references in the paper on topological invariants.

§ That is, to any set of arcs forming a circuit in one graph corresponds a set of arcs forming a circuit in the other.

III, Theorem 2 is slightly strengthened in the following corollary.

COROLLARY. *If G and G' satisfy the conditions of the above theorem and one of them is triply connected, then both are, and the two graphs are strictly isomorphic.*

APPLICATION TO ELECTRICAL THEORY. Let us say two graphs are electrically equivalent if there is a $1-1$ correspondence between their arcs so that if corresponding arcs of the two graphs are replaced by the same arbitrary electrical elements, then the same current will flow through corresponding elements. As R. M. Foster has stated (see reference in the last paper), two 2-isomorphic graphs are electrically equivalent. Is the converse true? If not, then it follows immediately from the above theorem that in some two electrically equivalent graphs, we can find a circuit in one corresponding to a subgraph of nullity 0 in the other. If we replace the elements of the circuit by a cell and conductors, a current will flow. But no current can flow through a network of nullity 0. Hence the two graphs are not electrically equivalent. Thus *two graphs are electrically equivalent if and only if they are 2-isomorphic, or, if and only if there is a $1-1$ correspondence between their arcs so that circuits correspond to circuits.*

2. The remainder of the paper is devoted to proving the theorem. We prove first a lemma. If X is a subgraph of G, we shall always denote by X' the corresponding subgraph of G'. We let a subgraph contain only those vertices which are on arcs of the subgraph.

LEMMA 1. *If G and G' satisfy the conditions of the theorem and H is a non-separable subgraph of G, then H' is non-separable.*

It is easily seen that if circuits correspond to circuits in two graphs, then subgraphs of rank i, nullity j, correspond to similar subgraphs. We need merely build up the two subgraphs arc by arc, and note that when corresponding arcs are added, the nullity of one graph increases if and only if the nullity of the other does.

If the lemma is false, then we can put $H' = I'_1 \dotplus I'_2$, where I'_1 and I'_2 each contain an arc, and $R(H')$* $= R(I'_1) + R(I'_2)$ (I, Theorem 13). Hence, by the above remark, $R(H) = R(I_1) + R(I_2)$, contradicting I, Theorem 14.

Suppose now we have proved the theorem for the case that both graphs are non-separable; then it follows for the general case. For if one of the

* That is, the rank of H'.

graphs, say G, were separable, let $G_1, \cdots, G_m$ be its components. From the lemma it follows that the corresponding subgraphs $G'_1, \cdots, G'_m$ of G' are the components of G'. The conditions of the theorem are satisfied for each pair of graphs G_i, G'_i $(i = 1, \cdots, m)$, hence G_i and G'_i are strictly 2-isomorphic. Having altered each G_i by operations of type (3) so that it becomes strictly isomorphic with G'_i, we bring G into strict isomorphism with G' by operations of type (2).

3. We now assume that both G and G' are non-separable. By I, Theorem 19,* we can build up G in the following manner. Take first an arc H_0. Next add an arc or chain, which with H_0 forms a circuit H_1 (if G was not H_0). Next add an arc or suspended chain, forming with H_1 the non-separable graph H_2 (if G was not H_1), etc. The subgraph H_0 of G, being a single arc, is strictly isomorphic with H'_0 (if $G = H_0$ is a 1-circuit, so is G'). If H_1 is not strictly isomorphic with H'_1, we alter G by a number of operations of type (3) so that it becomes so. If now H_2 is not strictly isomorphic with H'_2, we alter G again, etc. Thus to prove the theorem, we need merely show that *if H is a non-separable subgraph of G strictly isomorphic with H' and A is a chain in G with just its two end vertices in H, then we can alter G by operations of type* (3) *so that $K = H \dot{+} A$ becomes strictly isomorphic with K'.*

4. The graph H, together with the arc or arcs of A, form a circuit of graphs M_1. The corresponding (non-separable) subgraphs of G' form a circuit of graphs M'_1. For, K being non-separable, so is K', by Lemma 1; hence the graphs in M'_1 are not the components of H', and some subset of them form a circuit of graphs (I, Theorem 17). But no proper subset of them do, for the resulting graph would be non-separable (I, Theorem 16), while the corresponding graph in G is not; therefore the whole set forms a circuit of graphs.

H is by hypothesis strictly isomorphic with H', while K is not strictly isomorphic with K'. Group the arcs of A with the graph H into connected

---

* This theorem, and with it, I, Theorem 18, can be proved most simply as follows. G contains a circuit H_1 (I, Theorems 8 and 4). If $G \neq H_1$, let a be a vertex not in H_1, and let A be a chain from a to a vertex b of H_1. By I, Theorem 6, there is a chain B from a to a vertex c of H_1 not passing through b. From A and B we pick out a chain C with just its two end vertices b and c in H_1. Put $H_2 = H_1 \dot{+} C$. If $G \neq H_2$, we find similarly a chain forming with H_2 a non-separable graph H_3, etc. We arrive finally at a graph $H_s = G$.

subgraphs $G_1, G_2, \cdots, G_n$, such that each G_i is strictly isomorphic with G'_i (hence each G'_i is connected), while no two of these graphs together have this property. The graphs $G_1, \cdots, G_n$ satisfy the conditions for being a circuit of graphs, except that some of them may be separable (they are at least connected). We shall say they form a *generalized circuit of graphs* M_2. Of course $G'_1, \cdots, G'_n$ also form a generalized circuit of graphs M'_2.

5. The proof of the theorem rests on the following two facts.

(A) *Suppose two graphs G_j and G_k have no common vertex, while G'_j and G'_k have a common vertex. Then we can alter G by operations of type* (3) *so that G_j and G_k will get a common vertex. The graphs $G_1, \cdots, G_n$ form at each step a generalized circuit of graphs.*

(B) *Suppose G_j and G_k, also G'_j and G'_k, have a common vertex.* (1) *If there are at least three graphs in M_2, let a' and b' be the vertices of G'_k joining it to G'_j and some other graph of M'_2 respectively, and let a and b be the corresponding vertices of G_k (determined by the isomorphism between G_k and G'_k). Then if a is not the vertex of G_k joining it to G_j, we can alter G by an operation of type* (3) *so that it will become so.* (2) *If M_2 contains G_j and G_k alone and $G_j \dotplus G_k$ is not strictly isomorphic with $G'_j \dotplus G'_k$, then we can alter G by an operation of type* (3) *so that this will be true.*

With (A) and (B) proved, we prove the theorem as follows. If the supposition of (A) holds, we bring G_k into contact with G_j. If now the supposition of (B) holds, we turn G_k around so as to touch G_j correctly. Turning G_j around also if necessary, we bring $G_j \dotplus G_k$ into strict isomorphism with $G'_j \dotplus G'_k$. We employ (B) until any two graphs which touch, touch correctly. Now employ (A) again if necessary, etc. K is finally brought into strict isomorphism with K', as required.

6. We prove a lemma.

LEMMA 2. *If a and b are two vertices of G, and A_1, A_2, A_3 are three chains in G joining a and b and having no other common vertices, then A'_1, A'_2, A'_3 are chains, joining two vertices of G' and having no other common vertices.*

As $A_1 \dotplus A_2$, $A_1 \dotplus A_3$ and $A_2 \dotplus A_3$ are circuits, so are $A'_1 \dotplus A'_2$, $A'_1 \dotplus A'_3$, $A'_2 \dotplus A'_3$. Hence each A'_i is either a chain or a set of chains. But in the latter case, there are circuits in $A'_1 \dotplus A'_2 \dotplus A'_3$ besides those

named, which cannot be, as there are no such circuits in $A_1 \dot{+} A_2 \dot{+} A_3$ (see the proof of III, Theorem 2).

7. *Proof of* (A). Suppose the graphs $G_1, G_2, \cdots, G_n$ are named so that they lie in that order in M_2. Let $a_{k+1/2}$ be the vertex joining G_k and G_{k+1}, $(k = 1, \cdots, n)$ (putting $n + 1 = 1$). Suppose G_1 and G_i have no common vertex, while G'_1 and G'_i have a common vertex; we shall bring G_1 and G_i into contact.

The first step is to divide the graphs $G_2, \cdots, G_{i-1}$, and the graphs $G_{i+1}, \cdots, G_n$, into groups as follows. Suppose there is a chain C in G with just its two end vertices x and y in K. Suppose x lies in the graph G_p, and is not the vertex $a_{p+1/2}$, and y lies in the graph G_q, and is not the vertex $a_{q-1/2}$, and $1 < p < q < i$. Then we put the graphs $G_p, G_{p+1}, \cdots, G_q$ into the same group. If G_p and G_q, also G_q and G_r, fall into the same group, then we put all these graphs into the same group. Similarly for the graphs $G_{i+1}, \cdots, G_n$. Each group of graphs forms a subgraph of K. Let $I_2, \cdots, I_{j-1}$ be those formed from $G_2, \cdots, G_{i-1}$, and $I_{j+1}, \cdots, I_m$, those formed from $G_{i+1}, \cdots, G_n$. Put $I_1 = G_1$ and $I_j = G_i$. Then $I_1, I_2, \cdots, I_m$ *form a generalized circuit of graphs* M; say they lie in that order in M.

7'. Each I_k, formed from the graphs $G_{h_k}, G_{h_k+1}, \cdots, G_{h_{k+1}-1}$, say, is a *chain of graphs,* in that each graph G_p of the set is connected and has exactly one vertex in common with G_{p-1} and G_{p+1} (if they are in the set), and no other two have a common vertex. We shall show now that *the corresponding graphs of* M'_2 *form a chain of graphs.* This is trivial if there is just one graph in the set, which happens in particular if $k = 1$ or j. Suppose $1 \neq k \neq j$, and $h_{k+1} - h_k > 1$. The graph G_{h_k} is joined to some other graph G_{q_1} $(h_k < q_1 < h_{k+1})$ in I_k by a chain C_1 with only its end vertices x_1 and y_1 in K, where $x_1 \neq a_{h_k+1/2}$, $y_1 \neq a_{q_1-1/2}$. We shall form now three chains A_1, A_2, A_3, which have just two common vertices $c_1 \neq a_{h_k+1/2}$ and $d_1 \neq a_{q_1-1/2}$, lying in G_{h_k} and G_{q_1} respectively, and so that A_1 contains C_1 and possibly arcs of G_{h_k} or G_{q_1}, A_2 lies in $G_{h_k} \dot{+} \cdots \dot{+} G_{q_1}$, and $A_2 \dot{+} A_3$ is a circuit running around M.

If neither x_1 nor y_1 lies in H, C_1 and any circuit P running around K give such chains; then $c_1 = x_1$, $d_1 = y_1$. Suppose x_1 say, is in H; we must take care then to make $c_1 \neq a_{h_k+1/2}$. Join x_1 to $a_{h_k+1/2}$ by a chain D_1 in H. As $a_{h_k+1/2}$ is not a cut vertex of H, this graph being non-separable, there is a chain D_2 from $a_{h_k-1/2}$ to x_1 in G_{h_k} (if $a_{h_k-1/2}$ is not x_1) which does not pass through $a_{h_k+1/2}$ (see I, Theorem 6). Let c_1 be the first vertex of this chain on D_1; if D_1 contains $a_{h_k-1/2}$, let this vertex be c_1. Then A_1 is C_1 plus that

much of D_1 (if there is any) between x_1 and c_1, and A_2 and A_3 are two chains of a circuit P, which consists of the chain we have constructed in G_{h_k} joining $a_{h_k-1/2}$ and $a_{h_k+1/2}$, and a chain joining these vertices which runs around the other graphs of M_2. $d_1 = y_1$ as before.

Now by Lemma 2, A'_2 is a chain. It lies wholly in $G'_{h_k} \dotplus \cdots \dotplus G'_{q_1}$, and contains arcs of each of these graphs, as these statements are true when primes are dropped. Following from one end of A'_2 to the other, we pass through all these graphs, and hence they form a chain of graphs (remembering that $G'_1, \cdots, G'_n$ form a circuit of graphs). (We do not know in what order the graphs lie in the chain.)

If G_{q_1} is not $G_{h_{k+1}-1}$, there is a chain C_2 with just its two end vertices x_2 and y_2 in K, where $x_2 \neq a_{p_2+1/2}$ lies in a graph G_{p_2}, $y_2 \neq a_{q_2-1/2}$ lies in G_{q_2}, and $h_k \leqq p_2 \leqq q_1 < q_2 < h_{k+1}$. Hence the graphs $G'_{p_2}, \cdots, G'_{q_2}$ form a chain of graphs. It follows that the graphs $G'_{h_k}, \cdots, G'_{q_2}$ form a chain of graphs. Continuing in this manner, we see finally that $G'_{h_k}, \cdots, G'_{h_{k+1}-1}$ form a chain of graphs, as stated. From this fact we see at once that *$I'_1, \cdots, I'_m$ form a generalized circuit of graphs M'.*

8. Let $b_{k+1/2}$ be the vertex joining I_k and I_{k+1}, $k = 1, \cdots, m (m + 1 = 1)$. Put also $\alpha_k = I_k$ and $\alpha_{k+1/2} = b_{k+1/2}$, $k = 1, \cdots, m$. Suppose C is a chain in G with only its end vertices in K, joining b_k to b_l, or an inner vertex* of I_k to b_l, or an inner vertex of I_k to an inner vertex of I_l. Then we call the chain a (b_k, b_l), or an (I_k, b_l), or an (I_k, I_l) chain respectively. In any case, we can call it an (α_k, α_l) chain.

We now study what types of chains are possible in G. By symmetry, all the properties given below hold if the graphs I_1 and I_j, or the sets of graphs $\{I_p\}$ and $\{I_q\}$, $1 < p < j$, $j < q \leqq m$, are interchanged.

(a) *There is no (α_k, α_l) chain, such that $1 < k < l < j$, and for some integer p, $k < p + 1/2 < l$* (i. e. some vertex $b_{p+1/2}$ lies between α_k and α_l). This follows immediately from the definition of the graphs I_k.

(b) *There is no (I_1, I_j) chain.* For suppose there were. Then we can construct chains A_1, A_2 and A_3, with just the vertices c and d in common lying within I_1 and I_j respectively, and such that A_1 contains (I_1, I_j) and possibly arcs of I_1 or I_j, and $A_2 \dotplus A_3$ is a circuit going around M. (We use the proof in § 7′, with but slight changes). By Lemma 2, A'_2 and A'_3 are

---

* That is, a vertex of I_k which is neither $b_{k-1/2}$ nor $b_{k+1/2}$. We shall say such a vertex lies *within* I_k.

chains in K' having exactly two common vertices c', d'. I'_1 and I'_j each contain arcs of both A'_2 and A'_3, as this is true if primes are dropped; hence I'_1 and I'_j each contain one of the vertices c', d'. By hypothesis, $I'_1 = G'_1$ and $I'_j = G'_i$ have a common vertex; consequently one of the chains A'_2, A'_3 contains arcs of I'_1 and I'_j alone. But this cannot be, as A_2 and A_3 each contain arcs of graphs I_p, $p \neq 1, j$, and (b) is established.

(c) *There is no* (I_k, I_l) *chain,* $1 < k < j$, $j < l \leqq m$. For in this case, constructing chains A_1, A_2, A_3 as above, we find that the end vertices of A'_2 and A'_3 lie within I'_k and I'_l, and one chain passes through I'_1 while the other passes through I'_j; thus I'_1 and I'_j cannot have a common vertex, a contradiction.

(d) *There are no two chains* (I_1, α_k), (I_1, α_l), *with* $2 \leqq k < j$, $j < l \leqq m$. For suppose there were. Then following the proof in § 7′, we see that the graphs $I'_1, \cdots, I'_{k'}$ (k' is the greatest integer $\leqq k$) form a chain of graphs; similarly, $I'_{l'}, \cdots, I'_m, I'_1$ (l' is the smallest integer $\geqq l$) form a chain of graphs. Hence I'_1 touches both a graph I'_p, $1 < p \leqq k'$, and a graph I'_q, $l' \leqq q \leqq m$, and thus does not touch I'_j, a contradiction.

(e) *There are no two chains* $(\alpha_{k_1}, \alpha_{l_1})$, $(\alpha_{k_2}, \alpha_{l_2})$, *where* $1 \leqq k_1 \leqq j - 1$, $2 \leqq k_2 \leqq j$, $j < l_1 \leqq m + 1/2$, $j < l_2 \leqq m + 1/2$, *and* $k_1 < k_2$, *and for some integer* q, $l_1 \leqq q \leqq l_2$. For suppose there were. Consider first Case I: either $k_1 > 1$ or $k_2 < j$, say the former. Using $(\alpha_{k_1}, \alpha_{l_1})$, we form chains A_1, A_2 and A_3, and using $(\alpha_{k_2}, \alpha_{l_2})$, we form chains B_1, B_2 and B_3, as before, where A_2 and B_2 do not pass through I_1. As A'_2 and B'_2 are chains, the graphs $I'_{k'_1}, \cdots, I'_{l'_1}$, and also the graphs $I'_{k'_2}, \cdots, I'_{l'_2}$, form chains of graphs S'_1 and S'_2 respectively, neither of which contains I'_1, where k'_s is the smallest integer $\geqq k_s$, and l'_s is the largest integer $\leqq l_s$, $s = 1, 2$. (see § 7′).

Suppose first $k'_2 > k'_1$, $l'_2 > l'_1$. Starting at $I'_{k'_1}$, which is in S'_1 but not in S'_2, pass along the graphs of S'_1 towards I'_j, which is in both S'_1 and S'_2. As $I'_{l'_2}$ is in S'_2 but not in S'_1, we have not yet passed through all the graphs of S'_2; hence we can continue from I'_j into more graphs of S'_2. This shows that I'_j touches two graphs of the sets S'_1, S'_2, and thus does not touch I'_1, a contradiction.

Suppose next $k'_2 = k'_1$, $l'_2 > l'_1$. Then $k_2 = k'_2$, an integer, and the chain $(\alpha_{k_2}, \alpha_{l_2})$ has an inner vertex of I_{k_2} as end vertex. Hence B_2 has an end vertex within I_{k_2} (we can arrange that it has, as in § 7′); it follows that B'_2 has an end vertex within I'_{k_2}, as in (b). Therefore I'_{k_2} is at one end of the chain of graphs S'_2. I'_{k_2} (which is not I'_j) lies in S'_1, so we can follow S'_1 towards I'_j. We have not yet passed into $I'_{l'_2}$, which lies in S'_2 but not

in S'_1, so we can follow S'_2 further out of I'_j. Again I'_j has no vertex in common with I'_1, a contradiction.

The case $k'_2 > k'_1$, $l'_2 = l'_1$ is similar (this time $l_1 = l'_1 = q$, an integer). Suppose $k'_2 = k'_1$, $l'_2 = l'_1$. Then k_2 and l_1 are both integers, and $(\alpha_{k_2}, \alpha_{l_2})$ and $(\alpha_{k_1}, \alpha_{l_1})$ have vertices within I_{k_2} and I_{l_1} respectively. Hence I'_{k_2} is at one end of S'_2, and I'_{l_1} is at one end of S'_1. S'_1 and S'_2 contain the same graphs, and are thus the same chain; it has the distinct graphs I'_{k_2} and I'_{l_1} as ends, and thus has I'_j in its interior, so I'_j does not touch I'_1, a contradiction.

We have left Case II to consider, where $k_1 = 1$, $k_2 = j$. In this case $I'_j, \cdots, I'_{l'_2}$ and $I'_{l'_1}, \cdots, I'_m$, I'_1 form chains of graphs with I'_j and I'_1 as end graphs, and it is seen that I'_1 and I'_j have no common vertex. (e) is now proved.*

9. We return now to the proof of (A). Let us say that a chain (α_k, α_l) alternates with the vertices b_p, b_q, if the numbers k, l, p, q, are all distinct, and $\alpha_k, b_p, \alpha_l, b_q$ lie in that, or the reverse, cyclic order in M. The next step is to show that there are two vertices b_{p_1}, b_{q_1}, $1 < p_1 < j$, $j < q_1 \leqq m + 1/2$, such that

(α) no chain (α_k, α_l) alternates with these vertices,

(β) I_j is in contact with one of these, and

(γ) either I_j is in contact with the other also, or it can be brought into contact with it by an operation of type (3) on G.

If there is no (I_j, α_s) chain, $s \leqq j - 1$ or $s \geqq j + 1$, then $b_{j-1/2}$ and $b_{j+1/2}$ form such a pair of vertices. Suppose there is such a chain. By (b) and (d), either $1 < s \leqq j - 1$ or $s \geqq j + 1$ for all such chains; say the latter is true (the other case is similar). Let q_1 be the smallest number such that $q_1 - 1/2$ is an integer, and for any chain (I_j, α_s) or $(b_{j-1/2}, \alpha_s)$, $s \leqq q_1$. Put $p_1 = j - 1/2$; then b_{p_1} and b_{q_1} are the required vertices. To prove (α), we note first that no (I_j, α_k) chain alternates with these vertices. Take now any number l, $j < l < q_1$. There is no (α_l, α_k) chain, $q_1 < k \leqq m + 1/2$, by (a). Suppose there were a chain (α_l, α_k), with $1 \leqq k < p_1$. By the definition of q_1, there is either a chain (I_j, α_t) or $(b_{j-1/2}, \alpha_t)$, with $q_1 - 1/2 \leqq t \leqq q_1$;

* We can state properties (*b*), (*c*), (*d*) and (*e*) in the single sentence: *It is not true that there is a chain* (a_{k_1}, a_{l_1}), *and there is a chain* (a_{k_2}, a_{l_2}), *where* $1 \leqq k_s \leqq j$, $j \leqq l_s \leqq m + 1$ $(= 1)$, $s = 1, 2$, *and for some integers* p *and* q, $k_1 \leqq p \leqq k_2$, $l_1 \leqq q \leqq l_2$.

in either case, (e) is contradicted (put $k = k_1$, $l = l_1$, $j - 1/2$ or $j = k_2$, $t = l_2$).

(β) holds; we show that (γ) holds. Suppose I_j is in contact with b_{p_1}, but not with b_{q_1}. Let J_1 be that subgraph of G containing $I_{p_1+1/2}$ $(= I_j)$, $\cdots, I_{q_1-1/2}$, all vertices which can be joined to these graphs by chains not containing b_{p_1} or b_{q_1}, and all other arcs of G which touch these vertices. If J_2 is the complementary subgraph of G, then J_2 contains the graphs I_s, $s < p_1$ and $s > q_1$, as no chain alternates with b_{p_1} and b_{q_1}, and J_1 and J_2 have only these two vertices in common. Turning J_1 around at the vertices b_{p_1} and b_{q_1} brings I_j into contact with b_{q_1}, as required.—To continue the analysis, using the same notation, we should now rename all graphs and vertices α_k for $p_1 < k < q_1$, replacing subscripts k by $p_1 + q_1 - k$, so that the renamed graphs lie in cyclic order in K. I_j is then renamed $I_{p_1+q_1-j}$; it is this graph we must bring into contact with I_1.

9′. If $p_1 = 1 + 1/2$ or $q_1 = m + 1/2$, I_j is brought into contact with I_1. Suppose not. Then we shall find another pair of vertices b_{p_2} and b_{q_2} with the properties (α), (β) and (γ)—perhaps after we have performed the operation of type (3) described above—, that are nearer to I_1 in that $p_2 \leqq p_1$, $q_2 \geqq q_1$, and either $p_2 < p_1$ or $q_2 > q_1$. If one of these is $b_{1+1/2}$ or $b_{m+1/2}$, I_j can be brought into contact with I_1; if not, we find another such pair of vertices b_{p_3}, b_{q_3}, still nearer to I_1. After a finite number of steps we bring I_j into contact with I_1.

If no chain alternates with the vertices b_{p_1-1} and b_{q_1}, we can put $p_2 = p_1 - 1$, $q_2 = q_1$. I_j can be brought into contact with b_{q_2} if it is not already in contact; (α), (β) and (γ) now hold. Suppose there is a chain alternating with b_{p_1-1} and b_{q_1}; such a chain must be of the form (α_s, α_t), with $p_1 - 1/2 \leqq s \leqq p_1$, and $t > q_1$ or $t = 1$. We shall show then that no chain alternates with b_{p_1} and b_{q_1+1}, from which follows that we can put $p_2 = p_1$, $q_2 = q_1 + 1$. A chain alternating with these vertices must be of the form (α_l, α_k), with $q_1 \leqq l \leqq q_1 + 1/2$, $1 \leqq l < p_1$. But the chains (α_s, α_t) and (α_l, α_k) cannot both exist. For if $s = p_1$ and $t > q_1$, (e) is contradicted; if $s = p_1$ and $t = 1$, then $k = 1$ contradicts (d), and $k > 1$ contradicts (e). If $s = p_1 - 1/2$, then $t > q_1 + 1/2$ or $t = 1$, by (c); (d) is contradicted if $t = k = 1$, otherwise (e) is contradicted. (A) is now proved.

10. *Proof of* (B). We show first that b and a are the vertices joining G_k to G_j and another graph of M_2 respectively. If P' is a circuit running around M'_2, then that part of P' in G'_k is a chain A'_0 with end vertices a' and b'. P is a circuit; A_0 is those arcs of P lying in G_k, and is a chain with

end vertices a and b, as G_k is strictly isomorphic with G'_k. The statement now follows.

There is no chain in G joining a vertex of G_k to a vertex of the other graphs G_s which does not contain a or b. For suppose there were. We construct three chains A_1, A_2, A_3, as we have so often done, joining an inner vertex c of G_k to a vertex of K not in G_k. $A_2 \dotplus A_3$ is a circuit passing around M_2; say A_2 contains b, and A_3, a. First suppose A_2 contains arcs of G_j and G_k alone; then A'_3 contains arcs of G'_j and G'_k alone. As G_k is strictly isomorphic with G'_k and A_3 does not contain a, the arcs of A'_3 in G'_k do not contain a'. Thus A'_3 is not a chain, a contradiction. Suppose now A_2 contains arcs of other graphs besides; then A_3 contains no arcs of G_j. A_3 does not contain b; hence the arcs of A'_3 in G'_k do not contain b'. As A'_3 contains no arcs of G'_j, it is not a chain, a contradiction again.

Consequently, if we define J_1 containing G_k and J_2 containing the other graphs of M_2, as we did in proving (A), we can turn J_1, and with it, G_k, around at the vertices a and b, proving (B), (1). If the supposition of (B), (2) holds, the above operation can be performed, and $G_j \dotplus G_k$ is made strictly isomorphic with $G'_j \dotplus G'_k$. This completes the proof of the theorem.

PRINCETON UNIVERSITY,
HARVARD UNIVERSITY.

ON THE FUNDAMENTAL GROUP OF AN ALGEBRAIC CURVE.

By Egbert R. van Kampen.

The complex points of an algebraic curve of degree n in a complex projective plane form a 2-dimensional complex $\boldsymbol{C}$ (manifold but for the singular points) in a 4-dimensional manifold $\boldsymbol{P}$. The generators of the fundamental group of $\boldsymbol{P}-\boldsymbol{C}$ have been determined by Picard * and Lefschetz † as n loops in one of the lines of $\boldsymbol{P}$ round the n branches of $\boldsymbol{C}$. The relations have been determined implicitly by Enriques.‡ Zariski § pointed out that Enriques' results imply that a set of relations for these generators can be found on determining their transforms when the line containing them is moved round all singularities of the curve and round all the tangents from the origin to the curve. As the resulting proof seemed too algebraic for this simple and nearly purely topological question, Dr. Zariski asked me to publish a topological proof which is contained in this paper.¶ The method consists in cutting $\boldsymbol{P}$ so that it becomes a simpler space of which the group is readily found and then considering what happens to the group if the cuts are removed.

1. We introduce a kind of coördinate system in $\boldsymbol{P}$ by means of a point A not on $\boldsymbol{C}$ and a line α (= 2-dimensional manifold) not containing A. Among the lines through A there are only a finite number m having less than n points in common with $\boldsymbol{C}$. They are the lines through singular points of $\boldsymbol{C}$ or tangent to $\boldsymbol{C}$ and determine m points, $A_1, \cdots, A_m$ in α. We may suppose that $A_1, \cdots, A_m$ are not on the curve $\boldsymbol{C}$. The set of lines through a point B of α, excluding the line AB can be transformed topologically into an open interval of an x, y plane. Any point of $\boldsymbol{P}$ not on the line AB can now be determined

* *Théorie des fonctions algébriques de deux variables*, I, p. 86.

† *L'analyse situs et la géometrie algébrique*, p. 33.

‡ "Sulla costruzione delle funzioni algebriche possedenti una data curva di diramazione," *Annali di matematica*, Series 4, Vol. 1 (1923), pp. 185-198.

§ "On the problem of existence of algebraic functions of two variables possessing a given branch curve," *American Journal of Mathematics*, Vol. 51 (1929), pp. 305-328.

¶ Enriques proved that if we take n substitutions $S_1, \cdots, S_n$ on p objects, satisfying the relations (3) and (4), then there exists an algebraic function of degree p on the plane $\boldsymbol{P}$ having $\boldsymbol{C}$ as branch curve and whose branches permute according to the substitutions $S_1, \cdots, S_n$. The topological consequences, that can be drawn from this are, as Dr. Zariski pointed out to me, independent of the actual construction of the function: the construction of the Riemann manifold of the function is sufficient and quite easy. However this would only prove that (3) and (4) form all relations for the fundamental group of $\boldsymbol{P}-\boldsymbol{C}$ if the substitutions could be constructed so as to satisfy (3) and (4) and not any given relation group-theoretically independent of (3) and (4).

by its coördinates x, y and two more coördinates determining its projection from A on α.

We join the point B to $A_1, \cdots, A_m$ by means of m simple arcs $a_1, \cdots, a_m$ having only B in common. A point set formed by the line α, cut open along the arcs a_i, including both edges of the cuts but not including the points A_i and B, is called $\boldsymbol{T}$. We transform $\boldsymbol{T}$ topologically into a closed interval of a t, u plane from which $2m$ boundary points (corresponding to the points A_i and the point B counted m times) have been removed.

To the points of a point set $\boldsymbol{R}$, consisting of all points, except A, in the lines joining A to the points of $\boldsymbol{T}$,* we have now assigned 4 coördinates x, y, t, u, thus transforming $\boldsymbol{R}$ topologically into the interior and part of the boundary of an interval in the 4-space of those coördinates. Because $\boldsymbol{T}$ is simply connected and does not contain the projection of any branch point of $\boldsymbol{C}$, this curve appears in $\boldsymbol{R}$ as a set of n 2-dimensional manifolds no two of which have a point in common and each having one point corresponding to any one point of $\boldsymbol{T}$.

We can now prove that the point set $\boldsymbol{S} = \boldsymbol{R} - \boldsymbol{C}$ and any subset of $\boldsymbol{S}$, corresponding to an arc or a point in $\boldsymbol{T}$ has as its fundamental group the free group generated by n loops, round the n points of $\boldsymbol{C}$, in any one of the lines through A contained in the set.

For the subset (line through A) corresponding to a point of $\boldsymbol{T}$, that is for an open interval in the x, y plane, of which n points have been taken away, this is well known. Restricting ourselves to $\boldsymbol{S}$ itself, we have to prove that any closed curve in $\boldsymbol{S}$ can be deformed into a subset of a line, t, u constant, and that a 2-cell, whose boundary is already in that line, can itself be deformed into that line, both deformations leaving the origin O of the fundamental group fixed. We prove that any closed set $\boldsymbol{F}$ in $\boldsymbol{S}$ can be deformed into the line by alternating the following two steps:

a) The x and y coördinates are unchanged. An interval in the t, u plane, containing all points of $\boldsymbol{F}$ in its interior or on its boundary, is contracted homothetically, thus defining the change in the coördinates t and u of any point of $\boldsymbol{F}$. This process has to be stopped just before any point of $\boldsymbol{F}$ would reach a position on $\boldsymbol{C}$.

b) The t and u coördinates are unchanged. In every line t, u constant, we construct (according to the metric in the t, u, x, y interval) circles, all of the same radius δ, smaller than the minimum distance from points of $\boldsymbol{C}$ to points of $\boldsymbol{F}$, and these circles are enlarged continuously and at the same speed.

* The points of $\boldsymbol{P}$ on lines, through A and points of a_i, have of course been assigned two points in $\boldsymbol{R}$, because the points of a_i have been assigned two points in $\boldsymbol{T}$.

Any point of $\boldsymbol{F}$ is carried along on the boundary of the circle that may touch it during this process. The process ends as soon as two circles touch each other, or when a circle comes within a preassigned distance of the boundary of its x, y interval, or when a circle reaches O.

At the end of each step b), after step a) has been made at least once, $\boldsymbol{F}$ is transformed into a subset of a definite compact subset of $\boldsymbol{T}$ each point of which could be moved according to step a) over a certain distance with a positive minimum. Thus there exist a positive number ϵ, such that each time step a) is executed, the length of the moving interval can be shortened by at least ϵ or to zero. It follows that the two steps applied alternately will finally deform $\boldsymbol{F}$ into the line, t, u constant containing O.

Another way of constructing this fundamental group would start by proving that $\boldsymbol{S}$ is homeomorphic with the product of $\boldsymbol{T}$ and the subset of $\boldsymbol{S}$ corresponding to one point of $\boldsymbol{T}$.

2.* We will now construct the fundamental group of the space $\boldsymbol{S}_1$ consisting of all points of $\boldsymbol{P} - \boldsymbol{C} - A$ in lines through A and points of $\boldsymbol{T}_1$. Here $\boldsymbol{T}_1$ consists of the line α, cut open along the arcs $a_2, \cdots, a_m$ including both edges of each cut, but not including $A_1, \cdots, A_m$ or B. Starting from $\boldsymbol{S}$ we can construct $\boldsymbol{S}_1$ by identifying the two homeomorphic subsets $\boldsymbol{U}$ and $\boldsymbol{V}$ of $\boldsymbol{S}$ the result being the set $\boldsymbol{W}$ of $\boldsymbol{S}_1$ all three corresponding to the arc a_1.

We take as (free) generators of the fundamental group of $\boldsymbol{S}$ n loops $g_1, \cdots, g_n$ in a line contained in $\boldsymbol{U}$, through the point O in that line and in α.

In $\boldsymbol{S}_1$ we find as an additional generator a small loop h round the line AA_1 in α and as relations

$$g_i = h\phi_{i1}(g_1, \cdots, g_n)h^{-1}, \qquad (i = 1, \cdots, n), \tag{1}$$

expressing the transformation which the element g_i undergoes, when its containing line is moved round AA_1 along h. We shall prove that we have found all necessary generators and relations for $\boldsymbol{S}_1$.

LEMMA. *A continuous transformation of a complex into $\boldsymbol{S}_1$ can be deformed and the complex subdivided in such a way, that every cell of the subdivision can be considered as transformed first into $\boldsymbol{S}$ and then into $\boldsymbol{S}_1$ by means of the indentification of $\boldsymbol{U}$ and $\boldsymbol{V}$.*

Proof. As the deformation will not involve $\boldsymbol{W}$ itself we can construct it on $\boldsymbol{S}$ and we can suppose that $\boldsymbol{U}$ has for its projection on the t, u plane the edge $u = u_0$ of the interval. As the complex is a closed point set we can find a number 2ϵ such that for any point x_1, y_1, t_1, u_1 of the complex for which $| u_1 - u_0 | < 2\epsilon$, the segment $x = x_1, y = y_1, t = t_1, | u - u_0 | \leq 2\epsilon$ is

* The paper printed immediately after this one contains a treatment of the method used in this section from a more general standpoint.

completely contained in S. We move all points x, y, t, u of the complex for which $|u - u_0| \leq \epsilon$ into the edge $u = u_0$ and those for which $\epsilon \leq |u - u_0| \leq 2\epsilon$ a distance $2\epsilon - |u - u_0|$ along these segments, thus defining our deformation. The two closed subsets of the complex transformed originally into W and into points situated at a distance ϵ from W on the U side can be separated on the complex by means of a subdivision of the complex into cells of sufficiently small diameter. This subdivision satisfies our condition, because on the complex in the new position any arc passing through W from a point on one side of W to a point on the other side has to contain points of both closed subsets and accordingly at least one point of the subdivision; thus the lemma is proved.

After applying the lemma to an arbitrary element of the fundamental group of S_1 we join each vertex of the subdivision to O by means of an arc, that is not allowed to leave W after once entering it, and can now write the element as a product of elements that can be considered as lying in S. These last can be written as $f(g_i)$, $f(g_i)h$, $h^{-1}f(g_i)$, $h^{-1}f(g_i)h$ according as to whether they run in S from U to U, from U to V, from V to U, from V to V, so that the original element can be expressed in terms of the g_i and h.

Each relation between the generators of S_1 can be represented by a 2-cell to which we can suppose that the lemma has already been applied. The cell can be so changed that the boundary of each cell of the subdivision becomes an element of the fundamental group, expressed in terms of the generators. To prove this we join each vertex of the subdivision that is not already transformed in O to O by an arc that is not allowed to leave W after once entering it and replace the vertex plus a small neighborhood by a 2-cell, subdivided in the same way as the neighborhood, of which the center is transformed into O and a strip near the boundary into the original position of the neighborhood, while the rest is distributed along the arc: Each vertex is now transformed into O and each 1-cell is now transformed into an element of the fundamental group of S_1. This element can be considered as lying in S because after the preceding construction the 2-cells of the subdivision can still be considered as transformed into S. Hence it can be expressed in terms of the generators by a deformation in S.

We define a new transformation for a 2-cell. Its boundary and part of its interior is transformed like the boundary of the old 2-cell and its interior cut open along the 1-cell. The rest is a smaller 2-cell, which we cut in half by an arc joining the points which are transformed into the endpoints of the 1-cell. Both halves are now transformed into the 2-cell which is described by the 1-cell during its deformation in such a way that the transformation of the whole is continuous. Is this done for every 1-cell of the subdivision then the boundary of each 2-cell of the subdivision is an element of the

fundamental group expressed in the generators, so that the original relation can be compounded out of other relations each being represented by a 2-cell of $\boldsymbol{S}_1$, that can be considered at the same time as a 2-cell of $\boldsymbol{S}$.* The sum of the exponents with which h appears in each of these relations must be zero because the boundary has to finish in $\boldsymbol{U}(\boldsymbol{V})$ if it starts there. It follows that the element h can be eliminated by means of the relation (1) without disturbing the special property of the representation of the 2-cell because the relations (1) can be represented in $\boldsymbol{S}$. The remaining relation between the generators g_i alone is valid in $\boldsymbol{S}$ and thus identically satisfied. It follows that our original relation was a consequence of the relations (1) alone.

3. $\boldsymbol{R}_1$ is defined as $\boldsymbol{S}_1$ plus the points of $\boldsymbol{P} - \boldsymbol{C} - A$ in the line AA_1. In $\boldsymbol{R}_1$ h is equal to the identity, so that (1) becomes:

$$g_i = \phi_{i1}(g_1, \cdots, g_n). \tag{2}$$

No new generator is needed in $\boldsymbol{R}_1$. In fact the projection of any element of the fundamental group on α can be so deformed by an arbitrarily small deformation as not to contain A_1. But if the deformation is sufficiently small we can move the points of the arc itself along corresponding paths for instance in lines through B, transforming the element into an element of $\boldsymbol{S}_1$, that can be expressed in terms of the g_i and h.

The only new relations are $h = 1$ and its consequences. In fact any new relation can be expressed by means of a 2-cell some of whose points are transformed into points of the line AA_1. We may suppose that an arbitrarily small neighborhood of all those points is formed by a subcomplex L of a subdivision of the cell, of which the boundary consists of a number of closed curves of which the transforms in $\boldsymbol{R}_1$ all have a point B_1 in common and are as near to AA_1 as we please but do not contain any of its points. For every point X we construct a corresponding point X_1 in the same line through B but with its projection on α in B_1. For each point Y of the boundary of L we construct an arc that does not touch AA_1, is contained in a line through B, starts at Y, finishes at the corresponding point Y_1, that was just constructed, and finally, changes continuously, when we move the point Y along a component of the boundary, from being degenerated into a point at B_1, where we start, to being a loop round A_1, counted a certain number of times, when we come back to B_1. The arcs corresponding to one component of the boundary from a 2-cell, of which the boundary is: the component of the boundary, the corresponding set of points Y_1 and the loop round A_1, counted a certain number of times. If we take the neighborhood away from the 2-cell and replace it by the sum of those 2-cells plus the set of points X_1 and identify

* Compare Lemma 2 of the paper in this Journal: "On some lemmas in the theory of groups."

boundary points of those sets, we get a 2-cell with the same boundary as the old 2-cell, but from which a number of interior 2-cells have been removed, the boundary of each interior 2-cell being transformed into a loop round A_1 in α counted a certain number of times. By cutting the 2-cell open along arcs from the origin of the fundamental group to a point of each of those loops we change our relation into another relation, already valid in $\boldsymbol{S}_1$, and equivalent to the original relation, because the insertion of those loops round A_1 in α in the boundary of the 2-cell means the insertion of transforms of certain powers of h in the original relation and $h = 1$ in $\boldsymbol{R}_1$.

4. The process of the last two sections can be repeated for all points A_i giving rise to a total of m sets of relations of the form (2):

$$(3) \qquad g_i = \phi_{ij}(g_1, \cdots, g_n); \qquad (i = 1, \cdots, n; \ j = 1, \cdots, m),$$

for the point set formed by all points of $\boldsymbol{P} - \boldsymbol{C}$ except those in the line AB. But then we can repeat the reasoning of section 3, proving that for $\boldsymbol{P} - \boldsymbol{C}$ the only extra relation must express the fact, that a small loop round the line AB is equal to the identity. We find

$$(4) \qquad g_1 g_2 \cdots g_n = 1.$$

To formulate our result in the easiest way we take the origin of the fundamental group in A.

To determine the fundamental group of a projective plane $\boldsymbol{P}$ *minus an algebraic curve* $\boldsymbol{C}$, *take a point* A *not on* $\boldsymbol{C}$ *and a line* α *not containing* A. *Determine in* α m *points* A_i *by means of the lines through* A *and tangent to* $\boldsymbol{C}$ *or through singular points of* $\boldsymbol{C}$. *Take in a line through* A, *but not through any of the* A_i, n *loops* g_i *from* A *round the* n *points of* $\boldsymbol{C}$ *in that line, capable of generating the group of* $\boldsymbol{P} - \boldsymbol{C}$. *The relations between those elements are* (3) *and* (4). *The functions* $\phi_{ij}(g_1, \cdots, g_n)$ *represent the element into which* g_i *is transformed, when its containing line is moved, so that its intersection with* α *describes a loop round* A_j. *The* m *loops in* α *must be capable of generating the fundamental group of* $\alpha - \Sigma A_i$.

This last condition is necessary because in section 2 when we move the origin of the fundamental group into the arc a_i we have to do that along a path already contained in our space at that stage of construction.

It ought to be remarked that the line AB can be taken to be one of the lines AA_i. From our reasoning it follows then that one of the m sets of relations (3) is a consequence of all the others. Considering the space just before the line AB was added we find that in computing the fundamental group of a curve, degenerating into another curve and a line it is unnecessary to take in account the relations resulting from the intersection of the line with the rest of the curve.

THE JOHNS HOPKINS UNIVERSITY.

ON THE CONNECTION BETWEEN THE FUNDAMENTAL GROUPS OF SOME RELATED SPACES.

By EGBERT R. VAN KAMPEN.

Since the preceding paper contains the treatment of an example, not of a general theorem, its topological background does not appear very clearly. In this paper we treat the general theorem which underlies the contents of section 2 of the preceding paper. Other applications of this general theorem are to be found in the literature, for instance in a paper by K. Brauner,* but on the other hand the opportunity of simplifying the treatment of a fundamental group by means of this theorem has been overlooked several times, for instance in the same paper by Brauner and in a paper by W. Burau.† For this reason we did not think it superfluous to devote a separate paper to it.

The object of our theorem is to find the fundamental group of the space that results when certain homeomorphic subsets of a given space are identified. The first section gives the definition of that process of identification; the second section contains a lemma on the deformation of complexes; and the third section contains certain conditions and a lemma helping us over the point set-theoretic difficulties of the problem, while in the fourth section the fundamental group is constructed. In the fifth section we give the two special cases that will be most often useful. In the last, the path is shown to a more general theorem, of which however the general formulation would be more confusing than helpful, so that it is suppressed.

1. Suppose that a separable, regular, topological space $\boldsymbol{A}$ contains a subset $\boldsymbol{B}$, and a neighborhood $\boldsymbol{U}$ of $\boldsymbol{B}$, such that:

(a) $\boldsymbol{B}$ is closed and (b) $\boldsymbol{U}-\boldsymbol{B}$ is the sum of a finite or countable number of open sets $\boldsymbol{M}_i$, having no point in common. Then we can construct in a unique way a new separable, regular, topological space $\boldsymbol{C}=\boldsymbol{A}-\boldsymbol{U}+\Sigma\boldsymbol{N}_i$, where: (c) $\boldsymbol{N}_i$ is homeomorphic with $\boldsymbol{B}+\boldsymbol{M}_i$ (the set corresponding to $\boldsymbol{B}$ being called $\boldsymbol{B}_i$), is open in $\boldsymbol{C}$ and no two $\boldsymbol{N}_i$ have a point in common; (d) $\Sigma\boldsymbol{B}_i$ is closed in $\boldsymbol{C}$; (e) $\boldsymbol{C}-\Sigma\boldsymbol{B}_i$ is homeomorphic with $\boldsymbol{A}-\boldsymbol{B}$.

It follows from these properties that (f) the homeomorphisms of $\boldsymbol{C}-\Sigma\boldsymbol{B}_i$ with $\boldsymbol{A}-\boldsymbol{B}$ and of each $\boldsymbol{B}_i$ with $\boldsymbol{B}$ define a univalued continuous transformation T of $\boldsymbol{C}$ into $\boldsymbol{A}$. The $\boldsymbol{B}_i$ have the following properties: (g) They are

* "Zur geometrie der Funktionen zweier komplexen Variablen, IV," *Hamburger Abhandlungen*, Vol. 6 (1928), pp. 34-55.

† "Kennzeichnung der Schlauchknoten," *Hamburger Abhandlungen*, Vol. 9 (1932) pp. 125-133.

all homeomorphic with $\boldsymbol{B}$ and no two of them have a point in common; (h) Any sum of sets $\boldsymbol{B}_i$ is closed in $\boldsymbol{C}$.

On the other hand $\boldsymbol{A} = \boldsymbol{C} - \Sigma \boldsymbol{B}_i + \boldsymbol{B}$ is uniquely determined, if the sets $\boldsymbol{C}$ and $\boldsymbol{B}_i$ with the properties (g) and (h) are given and the properties (a) and (e) are required provided a condition is given to determine which (completely divergent) sequences of $\boldsymbol{C} - \Sigma \boldsymbol{B}_i$ correspond to sequences in $\boldsymbol{A} - \boldsymbol{B}$ convergent to a point of $\boldsymbol{B}$. If a metric of $\boldsymbol{C}$ is given, such that the distances of the pairs of points in the sets $\boldsymbol{B}_i$, corresponding to any pair of points in $\boldsymbol{B}$, have a positive lower limit, and that the distances of the pairs of sets $\boldsymbol{B}_i$ have a positive lower limit, this can be done by the following condition: (i) $\boldsymbol{A}$ sequence of points in $\boldsymbol{A} - \boldsymbol{B}$ converges to a point x of $\boldsymbol{B}$, provided that the distance between a point of the corresponding sequence in $\boldsymbol{C} - \Sigma \boldsymbol{B}_i$ and the nearest point x_i, corresponding to x, converges to zero.

Remarks. $\boldsymbol{C}$ is uniquely determined only if the actual subdivision of $\boldsymbol{U} - \boldsymbol{B}$ into the $\boldsymbol{M}_i$ is given. $\boldsymbol{A}$ is uniquely determined only if the actual homeomorphisms between the sets $\boldsymbol{B}_i$ and $\boldsymbol{B}$ are given.

A metric for $\boldsymbol{C}$ as used for the formulation of condition (i) can always be constructed. If the number of sets $\boldsymbol{B}_i$ is finite, condition (i) can be formulated independent of the metric of $\boldsymbol{C}$. If the number of sets $\boldsymbol{B}_i$ is infinite it is not possible to define $\boldsymbol{A}$ in a topologically invariant way, at least not if $\boldsymbol{A}$ must be regular. But a metric of $\boldsymbol{A}$ (which involves the regularity of $\boldsymbol{A}$) is used essentially in the proof of Lemma 2.

The first process might be called: "Cutting $\boldsymbol{A}$ along $\boldsymbol{B}$," and the second "Identifying the subsets $\boldsymbol{B}_i$ of $\boldsymbol{C}$." From section 3 on we will suppose that sets $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{B}_i, \boldsymbol{M}_i, \boldsymbol{N}_i$ are given satisfying the conditions introduced above.

Proof. If $\boldsymbol{A}$ is given and $\boldsymbol{C}$ can be constructed, the sets in $\boldsymbol{C}$ corresponding to open subsets of $\boldsymbol{A} - \boldsymbol{B}$ or of any one set $\boldsymbol{M}_i + \boldsymbol{B}$ have to be open in $\boldsymbol{C}$, and any open subset in $\boldsymbol{C}$ must be a sum of subsets of these types. But the set $\boldsymbol{A} - \boldsymbol{U} + \Sigma \boldsymbol{N}_i$ with those sets as neighborhoods has all the properties we assigned to $\boldsymbol{C}$. h) follows from the fact, that $\boldsymbol{C} - \Sigma \boldsymbol{B}_i$ and all the $\boldsymbol{N}_i$ are open in $\boldsymbol{C}$, so that the complement of any sum of $\boldsymbol{B}_i$'s is a sum of open sets in $\boldsymbol{C}$. The continuity of the transformation T follows from the fact that T is homeomorphic on certain open subsets of $\boldsymbol{C}$, covering $\boldsymbol{C}$.

If $\boldsymbol{C}$ is given and $\boldsymbol{A}$ can be determined, the set in $\boldsymbol{C}$ corresponding to any open set in $\boldsymbol{A}$ is an open set in $\boldsymbol{C}$ having corresponding point sets in common with all sets $\boldsymbol{B}_i$. As a consequence of condition (i), a point set is open in $\boldsymbol{A}$ provided it corresponds to an open set $\boldsymbol{U}$ in $\boldsymbol{C}$ having with all sets $\boldsymbol{B}_i$ corresponding point sets in common, and such that the distances of every

set of corresponding points x of $\boldsymbol{B}_i$ to the boundary of $\boldsymbol{U}$ have a positive lower limit. But the space defined by the set of points $\boldsymbol{C} - \Sigma \boldsymbol{B}_i + \boldsymbol{B}$ with all those open sets as neighborhoods has all the properties that we assigned to $\boldsymbol{A}$.

2. Lemma 1. *A given deformation of a subcomplex L of a singular complex K * can be extended to a deformation of K itself, leaving every point of K invariant that belongs to a simplex, which together with its boundary is contained in $K - L$. K can be an infinite complex,† if L contains all but a finite number of the simplexes of K.*

Proof. By the given deformation of L and our condition on the simplexes not touching L the deformation is defined for all vertices of K, so that an induction proof of the lemma can be completed by defining the deformation for the interior of any simplex S, provided it has already been defined on its boundary T. The simplex S plus the complex described by T during its deformation can be transformed topologically into a simplex R in such a way that S is transformed into a simplex R_1 homothetic with R and in its interior. The homothetic deformation of R_1 into R gives the extension of the deformation to S.

3. To provide a connection between the fundamental groups of $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ we need the following restrictions:

(1) In any neighborhood $\boldsymbol{V}$ of a point of $\boldsymbol{B}_i$ in $\boldsymbol{N}_i$ there is contained another neighborhood $\boldsymbol{V}'$, such that any singular 0, 1, 2 sphere in $\boldsymbol{V}'$ is the boundary of a singular 1, 2, 3 cell in $\boldsymbol{V}$. If the number of sets $\boldsymbol{B}_i$ is infinite it must be possible to take $\boldsymbol{V}'$ as the set in $\boldsymbol{N}_i$ corresponding to the intersection of $\boldsymbol{M}_i + \boldsymbol{B}$ and a certain open set in $\boldsymbol{A}$ if $\boldsymbol{V}$ is taken to be the set in $\boldsymbol{N}_i$ corresponding to the intersection of $\boldsymbol{M}_i + \boldsymbol{B}$ and a given open set in $\boldsymbol{A}$.

(2) In any neighborhood $\boldsymbol{W}$ of a point of $\boldsymbol{B}$ in $\boldsymbol{B}$ there exists another neighborhood $\boldsymbol{W}'$, such that any singular 0, 1 sphere in $\boldsymbol{W}'$ is the boundary of a singular 1, 2 sphere in $\boldsymbol{W}$.

Lemma 2. *Any complex K whose dimension is at most two, and that is transformed continuously into $\boldsymbol{A}$, can be deformed and subdivided in such a way, that the new transformation of each simplex of the subdivision into $\boldsymbol{A}$ can be written as the product of a transformation of that simplex into $\boldsymbol{C}$ and the transformation T of $\boldsymbol{C}$ into $\boldsymbol{B}$ defined in* (f). *We write for brevity: that each simplex of the subdivision has property T.*

---

* A complex transformed into a certain space by a univalued continuous transformation.

† S. Lefschetz, *Topology*, Ch. VII.

Proof. We suppose, that K is subdivided already in such a way, that any simplex of K touching $\boldsymbol{B}$ is completely contained in $\boldsymbol{U}$. We transform the part of K common to the sum of all those simplexes and $\boldsymbol{M}_i + \boldsymbol{B}$ into $\boldsymbol{N}_i$, calling the closed point set corresponding to the points of K in $\boldsymbol{B}$: L_i, and the rest of the corresponding point set in $\boldsymbol{N}_i$, after subdividing it into an infinite complex: K_i. The lemma follows if we succeed in constructing a deformation of the sum of all but a finite number of the simplexes of all complexes K_i, into $\boldsymbol{B}_i$ such that any sequence of points of K_i, converging to a point of L_i, continues to converge to that point during the process of deformation. In fact the deformations in the different $\boldsymbol{N}_i$ correspond to deformations in $\boldsymbol{M}_i + \boldsymbol{B}$ leaving the points of K in $\boldsymbol{B}$ fixed, so that they can be combined into one deformation; according to Lemma 1 this deformation can be extended to a deformation of K itself and after this deformation any simplex of K, that has a point in common with $\boldsymbol{B}$ but is not contained in $\boldsymbol{B}$, is contained in one of the sets $\boldsymbol{M}_i + \boldsymbol{B}$. From this the lemma follows.

The deformation of K_i can be constructed, following an example of S. Lefschetz,* by means of the metric of $\boldsymbol{N}_i$ induced by a metric of $\boldsymbol{A}$ and a short induction construction of which we only give the last step. In other words, we assume that the deformation has been constructed for the sum of all 1-simplexes of K_i, except a finite number of them. Because of the condition on convergent sequences, the compactness of $L_i + K_i$, and conditions (1) and (2) there exists a positive number ϵ, a function of the positive number δ, such that for any 2-simplex of K_i, of which the diameter is smaller than ϵ, there exists a 3-cell of diameter less than δ, whose boundary consists of: the 2-simplex of K_i, the complex described by its boundary simplexes during their deformation and another 2-simplex contained in $\boldsymbol{B}$. Because of the second part of (1) we can take ϵ to be the same function of δ for all K_i even if their number is not finite. There is at most a finite number of simplexes in all the complexes K_i together, of which the diameter is more than any positive number ϵ. We take a sequence of numbers δ_i, converging to zero, and take as deformation-cell of any 2-simplex of K_i, of which the diameter λ satisfies: $\epsilon_{i+1} \leq \lambda < \epsilon_i$ a 3-cell of diameter less than δ_i. It is clear that the deformation determined by all those deformation-cells satisfies our condition.

4. For the fundamental group of any space $\boldsymbol{P}$ we write: $\mathfrak{G}(\boldsymbol{P})$. In order to determine the fundamental group of $\boldsymbol{A}$ by means of properties of $\boldsymbol{B}$ and $\boldsymbol{C}$ we will assume the following conditions:

(3) $\boldsymbol{B}$ is arcwise connected;

---

* *Topology*, p. 93.

(4) $\boldsymbol{C}$ is the sum of a finite or countable number of arcwise connected components, each containing at least one set $\boldsymbol{B}_i$.

From (4) it follows immediately, that $\boldsymbol{A}$ is arcwise connected.

We call the components of $\boldsymbol{C}$: $\boldsymbol{C}_l$, $l = 1, 2, \cdots$; and one of the sets $\boldsymbol{B}_i$ contained in $\boldsymbol{C}_l$ we call $\boldsymbol{B}_l$, the others if they exist are called $\boldsymbol{B}_{lj}$, $j = 1, 2, \cdots$. As origin for $\mathfrak{G}(\boldsymbol{A})$ we take a point O in $\boldsymbol{B}$; the corresponding point in $\boldsymbol{B}_l$ is called O_l and is taken as origin of $\mathfrak{G}(\boldsymbol{C}_l)$; the corresponding points in $\boldsymbol{B}_{lj}$ are called O_{lj}. An arbitrary but fixed arc in $\boldsymbol{C}_l$ from O_l to O_{lj} is called h_{lj}. In $\boldsymbol{A}$ there is an element g_{lj} of $\mathfrak{G}(\boldsymbol{A})$ corresponding to each arc h_{lj}.

THEOREM 1. *As a complete set of generators of* $\mathfrak{G}(\boldsymbol{A})$ *we can take:*

(5) *A complete set of generators* $a_{l\alpha}$, $\alpha = 1, 2, \cdots$ *for each* $\boldsymbol{C}_l$;

(6) *All elements* g_{lj}.

Proof. We apply Lemma 2 to an arbitrary element g of $\mathfrak{G}(\boldsymbol{A})$. Each vertex of the subdivision can be deformed into the point O by a deformation of which the rest takes place entirely in $\boldsymbol{B}$ immediately after the vertex enters $\boldsymbol{B}$ for the first time. According to Lemma 1, this deformation can be extended to a deformation of the whole element g after which g can be written as a product of other elements, each having property T. If the transform of each factor in C, is a curve from O_{li} to O_{lj}, it can be written in the form $g_{li}^{-1}\phi(a_{l\alpha})g_{lj}$, which proves the theorem.

To a set of generators b_β, $\beta = 1, 2, \cdots$ of $\mathfrak{G}(\boldsymbol{B})$ there corresponds a set of generators $b_{l\beta}$ of $\mathfrak{G}(\boldsymbol{B}_l)$ that can be expressed in terms of the $a_{l\alpha}$:

$$(7) \qquad b_{l\beta} = \phi_{l\beta}(a_{l\alpha}),$$

and a set of generators $b_{lj\beta}$ of $\mathfrak{G}(\boldsymbol{B}_{lj})$. The elements of $\mathfrak{G}(\boldsymbol{C}_l)$, which may be written symbolically as $h_{lj}b_{lj\beta}h_{lj}^{-1}$ can also be expressed in terms of the $a_{l\alpha}$:

$$h_{lj}b_{lj\beta}h_{lj}^{-1} = \phi_{lj\beta}(a_{l\alpha}).$$

THEOREM 2. *As a complete set of relations for* $\mathfrak{G}(\boldsymbol{A})$ *we can take:*

(8) *A complete set of relations for each* $\mathfrak{G}(\boldsymbol{C}_l)$:

$$\psi_{lj}(a_{l\alpha}) = 1.$$

$$(9) \qquad \phi_{1\beta}(a_{1\alpha}) = \phi_{l\beta}(a_{l\alpha}) = g_{kj}^{-1}\phi_{kj\beta}(a_{k\alpha})g_{kj};$$

these last express the fact that the generators of $\boldsymbol{B}_i$ *(expressed in terms of* $a_{l\alpha}$ *and* g_{lj}*) are all equal in* $\boldsymbol{A}$.

Proof. A relation between the generators of $\mathfrak{G}(\boldsymbol{A})$ can be represented by a singular 2-cell in $\boldsymbol{A}$, of which the boundary is the set of generators, whose product is equal to the identity. On this 2-cell we apply Lemma 2. Each element of the boundary describes, during the deformation of Lemma 2, a 2-cell with property T. The deformed 2-cell plus all the 2-cells corresponding to the elements of its boundary can be joined together to form a new 2-cell P,

with the same boundary as the original one, but with a subdivision, each cell of which has property T and which does not subdivide the elements of the boundary. Applying the method of proof of the preceding theorem we can give a deformation in $\boldsymbol{B}$ of the sum of all 1-cells of P contained in $\boldsymbol{B}$, and then a deformation of the sum of all other 1-cells such that after the deformation each 1-cell originally contained in $\boldsymbol{B}$ is expressed in terms of the generators of $\mathfrak{G}(\boldsymbol{B})$, while all other 1-cells are expressed in terms of the generators of $\mathfrak{G}(\boldsymbol{A})$. Applying Lemma 1 we shall see that when this deformation has been extended to P itself, each cell of the subdivision still has property T, while the boundary of each cell is now an expression in the generators of $\mathfrak{G}(\boldsymbol{B})$ and $\mathfrak{G}(\boldsymbol{A})$. It follows that each relation between those generators is a consequence of other relations, each of which can be represented by a 2-cell with property T.*

Suppose that such a 2-cell Q with property T, can be transformed into $\boldsymbol{C}_l$ by a transformation S. If at least one generator g_{lj} occurs in the relation, then some vertex of Q is transformed by S into O_l and we can start reading our relation at that vertex.† If g_{lj} is the first element of this type that occurs then the succeeding elements in the relation are generators of $\boldsymbol{B}$, transformed by S into $\boldsymbol{B}_{lj}$ till finally we find the element g_{lj}^{-1}. As a result of (9) we can consider the elements transformed into $\boldsymbol{B}_{lj}$ as being of the form $g_{lj}^{-1}\phi_{lj\beta}(a_{l\alpha})g_{lj}$, so that without any deformation, the elements g_{lj} and g_{lj}^{-1} can be eliminated.

If the relation does not contain any element g_{lj}, or after all elements g_{lj} have been eliminated by the above process, each vertex of Q is transformed by S into the same point of $\boldsymbol{C}_l$. If this point is O_l, the relation is a relation in $\boldsymbol{C}_l$ between the generators of its fundamental group, and thus a consequence of (8); if this point is O_{lj}, all elements of the relation can be written in the form $g_{lj}^{-1}\phi_{lj\beta}(a_{l\alpha})g_{lj}$ and thus the relation is a relation valid in $\boldsymbol{C}_l$, but transformed by g_{lj} and thus again it is a consequence of (8). This proves our theorem.

5. The meaning of the two preceding theorems will be clearer after the formulation of the following two special cases. The first of these was used in the preceding paper, the other could have been used instead if the proof had been slightly altered.

COROLLARY 1. *We suppose that the number of sets* $\boldsymbol{B}$ *is two, and that* $\boldsymbol{C}$ *is arcwise connected. A set of generators for* $\mathfrak{G}(\boldsymbol{C})$ *is formed by the elements*

* Compare the lemmas in the succeeding paper.

† This refers to the fact that the relation represented by a 2-cell must always be considered as written cyclically. We can break the cycle open at any vertex and read the relation from that vertex.

a_a with $\psi_\beta(a_a) = 1$ as relations. A set of generators of $\mathfrak{G}(\boldsymbol{B})$ is b_γ, while $b_{i\gamma}$ $(i = 1, 2)$ are the corresponding elements in $\boldsymbol{B}_i$; h_i is an arc in C from the origin of $\mathfrak{G}(\boldsymbol{C})$ to the origin of $\mathfrak{G}(\boldsymbol{B}_i)$. The elements $h_i b_{i\gamma} h_i^{-1}$ of $\mathfrak{G}(\boldsymbol{C})$ can be expressed in terms of the a_a:

$$h_i b_{i\gamma} h_i^{-1} = \phi_{i\gamma}(a_a).$$

As generators for $\mathfrak{G}(\boldsymbol{A})$ we can take the elements a_a and another element g, and as relations:

$$\psi_\beta(a_a) = 1, \qquad g\phi_{1\gamma}(a_a) = \phi_{2\gamma}(a_a)g.$$

COROLLARY 2. *The number of sets $\boldsymbol{B}_i$ is two, and $\boldsymbol{C}$ is the sum of two arcwise connected components: $\boldsymbol{C}_1$, containing $\boldsymbol{B}_1$, and $\boldsymbol{C}_2$, containing $\boldsymbol{B}_2$. A set of generators for $\mathfrak{G}(\boldsymbol{C}_i)$ $(i = 1, 2)$ is formed by the elements a_{ia}, a set of relations by $\psi_{i\beta}(a_{ia}) = 1$. The elements $b_{i\gamma}$ corresponding in $\boldsymbol{B}_i$ to a set of generators b_γ of $\mathfrak{G}(\boldsymbol{B})$ can be expressed in terms of the a_{ia}. (If necessary after their beginnings and ends have been joined by the same arc to the origin of $\mathfrak{G}(\boldsymbol{C}_i)$): $b_{i\gamma} = \phi_{i\gamma}(a_{ia})$.*

As generators for $\mathfrak{G}(\boldsymbol{A})$ we can take the elements a_{1a} and a_{2a} and as relations:

$$\psi_{1\beta}(a_{1a}) = 1, \quad \psi_{2\beta}(a_{2a}) = 1, \quad \phi_{1\gamma}(a_{1a}) = \phi_{2\gamma}(a_{2a}).$$

6. By a repeated application of the preceding construction * the fundamental group of $\boldsymbol{A}$ can be found in the case where conditions (3) and (4) are replaced by the following:

$\boldsymbol{B}$ is the sum of a finite number of arcwise connected components,

$\boldsymbol{C}$ is the sum of a finite or countable number of arcwise connected components, each containing at least one of the sets corresponding to components of $\boldsymbol{B}$.

$\boldsymbol{A}$ is arcwise connected.

However the theorem can still be generalized, if the number of components of $\boldsymbol{B}$ is countable, provided that every sum of components of $\boldsymbol{B}$ is closed in $\boldsymbol{B}$.† For any element of the fundamental group of $\boldsymbol{A}$ or any 2-cell, representing a relation between these elements, is a compact set in $\boldsymbol{A}$, so that it can only meet a finite number of components of $\boldsymbol{B}$. It follows that each element and each relation of $\mathfrak{G}(\boldsymbol{A})$ has been taken in account after the process of identification has been performed a finite number of times.

THE JOHNS HOPKINS UNIVERSITY.

* We do not describe this process, as the resulting theorems are too complicated to be of much use, and the process can be readily set up for any special example.

† If this last condition is not verified $\mathfrak{G}(\boldsymbol{A})$ can still be constructed after the definition of a limit for elements of $\mathfrak{G}(\boldsymbol{A})$ as the closure of the group that will appear as a result of the continued identification process.

ON SOME LEMMAS IN THE THEORY OF GROUPS.

By Egbert R. van Kampen.

Group-theoretic constructions like the one in Corollary 2 of the preceding paper have been used by several authors as a tool for the theory of abstract groups.* In this paper we shall show how the lemmas, proved by these authors on group constructions of that type, can be proved by means of simple topological reasonings on 2-dimensional complexes in the plane. The resulting proofs for those lemmas are shorter and of much clearer construction than the original proofs, but nevertheless not essentially different. It is this 2-dimensional method of proof and not any original result which justifies the publication of this paper. The lemmas in section 1 of this paper give the connection between abstract groups and 2-cells, and are reminiscent of Dehn's theory of the "Gruppenbild." Their proof is only sketched. In section 2 the lemmas proved by the authors cited are restated in a generalized form in 3 theorems with 2 corollaries. Nowhere in this paper is a restriction placed on the number (power of infinity) of any set of generators or of relations used.

1. Lemma 1. *If a certain relation $W = 1$ between the elements $a_1, a_2, \cdots$ of a group $\mathfrak{G}$ is a consequence of the relations $R_1 = 1$, $R_2 = 1, \cdots$ (written cyclically in the shortest possible way) between those elements (that means if a product $\prod_1^n T_i R_{\nu_i}^{\epsilon_i} T_i^{-1}$, $\epsilon_i = \pm 1$, where T_i is a certain product in the elements a_i, can be reduced to W by simple contractions: $a_i a_i^{-1} = 1$), then there exists a plane, connected and simply connected complex $\mathbf{W}$ with the following properties:*

(a) *To every oriented 1-cell of the complex there is assigned one of the elements a_i or one of their inverses.*

(b) *If we start at a certain vertex of the boundary of any 2-cell of $\mathbf{W}$ and follow the boundary from that vertex in a certain direction, we shall find assigned to the 1-cells of that boundary the elements a_i, in the same order and with the same exponents as they occur in one of the relations $R_i = 1$.*

* O. Schreier, "Die Untergruppen der freien Gruppen," *Hamburger Abhandlungen*, Vol. 5 (1927), pp. 164-168; W. Magnus, "Über die diskontinuierlichen Gruppen mit einer definierenden Relation," *Crelle's Journal*, Vol. 163 (1930), pp. 141-165.

(c) *The same is true for the boundary of* **W** *itself as seen from the rest of the plane if we take the relation* $W=1$ *instead of one of the relations* $R_i=1$. [The vertex where we start reading this relation will be called O.]

Proof. If W is identically equal to $TR_i^{\epsilon_i}T^{-1}$, then the representing complex can be taken as one 2-cell, whose boundary has been divided in as many 1-cells as R_i contains elements, plus a segment, attached to the 2-cell at one of its vertices and subdivided into as many one cells as T contains elements. After assigning elements a_i to all the 1-cells, the proof of the theorem is immediate.

If W is identically equal to $\prod_1^n T_i R_{\nu_i}^{\epsilon_i} T_i^{-1}$, then a number $\Sigma\,|\,\epsilon_i\,|$ of complexes of the kind described above can be joined together at the endpoints of the segments to form the complex corresponding to this relation. Care must be taken that the order of the segments in the plane is the same as the order of the corresponding factors in the relation.

Now the lemma is proved in case the relation $W=1$ is simply $\prod_1^n T_i R_{\nu_i}^{\epsilon_i} T_i^{-1}$ without any contractions. But it is easy to extend the lemma from one relation to another that can be derived from it by a simple contraction. In fact the two successive elements a_i and a_i^{-1}, that are taken away may be represented on the original complex by either two successive 1-cells on its boundary or by one 1-cell of which one endpoint is not incident with another cell. In the first case the two 1-cells can be brought into coincidence by a deformation without any other change in the complex. In the second case the 1-simplex can be taken away. Clearly the result is, in both cases, the complex belonging to the new relation, which has all the properties described in the lemma.

LEMMA 2. *Suppose a connected and simply connected plane complex* **W** *is given and an element* a_i *or its inverse assigned to each oriented* 1-*cell of the complex. Then a product* W *of elements* a_i *can be assigned to the boundary of* **W** *as seen from the rest of the plane, by arranging the elements assigned to* 1-*cells of that boundary in the order and with the exponents as they occur there. In the same way a product* R_i *can be assigned to the boundary of each* 2-*cell of* **W**. *The product* W *is equal to the identity, provided that the products* R_i *are equal to the identity.*

Proof. We can suppose that no 1-cell of **W** has a free endpoint because any such 1-cell could be taken away from **W** without essential change in W. We assume that the theorem has been proved for all complexes of which the

number of 2-cells is less than n and prove it for a certain complex $\boldsymbol{W}$ with n 2-cells. Take the interior of a certain 2-cell P away from $\boldsymbol{W}$ and make the resulting complex again simply connected by cutting it open along a simple arc consisting of 1-cells of $\boldsymbol{W}$ from a point of the boundary of $\boldsymbol{W}$ to a point of the boundary of P (This arc may degenerate into a point). The product W is equal to the identity as a consequence of the relations represented by P and by the new complex. As the theorem is true if $\boldsymbol{W}$ has only one 2-cell it can be proved for any complex $\boldsymbol{W}$.

LEMMA 3. *The complex $\boldsymbol{W}$ corresponding to a relation $W = 1$, is the sum of a number of 2-dimensional elements * having isolated boundary points in common and a number of 1-dimensional complexes having isolated points in common with 1 or more of the 2-dimensional elements. By a succession of simple contractions in the relation $W = 1$ any 1-cells with free ends can be eliminated from $\boldsymbol{W}$. The relation $W = 1$ is a consequence of the relations corresponding to the 2-dimensional elements of $\boldsymbol{W}$.*

The first two parts follow immediately from the proof of Lemma 1, the last is a consequence of Lemma 2.

2. Suppose that a group $\mathfrak{G}$ is given by certain sets of generators, each set being represented by one letter: $b, a_1, a_2, a_3, \cdots$ and by certain sets of relations, each set being represented by one equation and only containing elements of the sets of generators written in the equation:

$$(1) \qquad R_1(a_1, b) = 1, \qquad R_2(a_2, b) = 1, \cdots.$$

THEOREM 1. *Any element p of this group may be represented as a product of factors, each of which contains elements of the set b and of one of the sets a_i. Suppose this has been done in two ways:*

$$(2) \qquad p = \prod_1^n \phi_i(a_{\nu_i}, b) = \prod_1^m \psi_i(a_{\mu_i}, b).$$

Then either $m = n$, $\mu_i = \nu_i$ and

$$(3) \qquad \phi_i(a_{\nu_i}, b) = T_i(b)\psi_i(a_{\mu_i}, b)T^{-1}_{i+1}(b), \qquad T_1(b) = T_{n+1}(b) = 1$$

or in at least one of the above products simplification can be effected because at least one factor can be expressed in the elements b alone or because two successive factors contain the same set of generators a_i.

Proof. The theorem follows from the special case, where $p = 1$, by applying it to the relation:

* An element is a complex homeomorphic with a cell.

$$\prod_1^n \phi_i(a_{\nu_i}, b) \prod_0^{m-1} \psi_{m-i}^{-1}(a_{\mu_{m-i}}, b) = 1.$$

To prove the special case we apply Lemma 1 to the relation

$$W = \prod_1^n \phi_i(a_{\nu_i}, b) = 1$$

in which we may suppose that no simple contractions are possible. Furthermore we may suppose, that the complex $\boldsymbol{W}$, representing $W = 1$, is a 2-dimensional element. For if this is not true then $\boldsymbol{W}$ contains at least one such element $\boldsymbol{E}$, that is only connected with the rest of $\boldsymbol{W}$ at one boundary point P and such that the point O (Lemma 1, c) is either not on $\boldsymbol{E}$ or in P. We can now prove the theorem for the relation represented by $\boldsymbol{W}$ if it has been proved in the case of a 2-dimensional element by using it for $\boldsymbol{E}$ and so reducing the number of 2-dimensional elements in $\boldsymbol{W}$.

We now suppose that $\boldsymbol{W}$ is a 2-dimensional element. If all 1-cells (including their endpoints) that correspond to elements b are taken away from $\boldsymbol{W}$ it will be divided into a certain number of components $\boldsymbol{C}_i$. *All 1-cells in each component $\boldsymbol{C}_i$ correspond to elements of the same set a_i.* Otherwise some 2-cells would have to contain 1-cells assigned to generators of two different sets a_i and a_j and this is impossible because no given relation contains generators from both sets a_i and a_j.

If $n = 1$ in (4) the theorem is trivial; if $n = 2$ and $\nu_1 \neq \nu_2$ the sum of all components $\boldsymbol{C}_i$ containing elements a_{ν_1} of $\phi_1(a_{\nu_1}, b)$ together with their boundaries and all 1-cells corresponding to elements b of $\phi_1(a_{\nu_1}, b)$ is a complex of which the boundary as seen from the rest of the plane corresponds to the factor $\phi(a_{\nu_1}, b)$ and a product of elements b. From Lemma 2 it follows that $\phi_1(a_{\nu_1}, b) = T(b)$ and thus that $\phi_2(a_{\nu_2}, b) = T_b^{-1}$.

In the general case we suppose that no two successive numbers ν_i are equal and then we have to prove that at least one of the factors is equal to a product of elements b. We take all components $\boldsymbol{C}_i$ of $\boldsymbol{W}$ containing 1-cells corresponding to elements a_{ν_i} together with their boundaries and the 1-cells of $\boldsymbol{W}$ corresponding to elements b in factors $\phi_l(a_{\nu_l}, b)$ for which $\nu_i = \nu_l$.

If this sum does not contain a component meeting cells corresponding to elements of two different factors $\phi_i(a_{\nu_i}, b)$, then all factors $\phi_l(a_{\nu_l}, b)$, with $\nu_l = \nu_i$, are equal to a product of elements b.

If this sum does contain a component meeting cells corresponding to elements of two different factors $\phi_i(a_{\nu_i}, b)$ and if these factors are taken as near to each other as possible in the given product, then the product of all factors in between them is equal to a product of elements b and contains a

fewer number of factors $\phi_i(a_{\nu_i}, b)$. It follows that Theorem 1 can be proved by means of an induction proof on the number of factors ϕ_i in W.

THEOREM 2. *Any relation valid in the group* $\mathfrak{G}$ *between elements* b *alone is a consequence of relations of the following type:*

$$T(b) = \prod_{l=1}^{m} \phi_l(a_i, b) T_l(b) \phi_l^{-1}(a_i, b), \tag{5}$$

each following from the set of relations $R_i = 1$ *alone. At the moment that each relation is used, all factors* $T_l(b)$ *must be known to be equal to the identity as a consequence of other relations of the type* (5) *or of only one set of relations* $R_i = 1$.

COROLLARY 1. *If the same set of relations between the elements* b *follow from any of the sets of relations* $R_i = 1$ *separately, then no more such relations follow from all relations* $R_i = 1$ *together.**

Then the first relation of the type (5) used to find an additional relation $T(b) = 1$ would already give that relation as a consequence of the set of relations $R_i = 1$ alone.

Proof. Take any one of the components $\boldsymbol{C}_i$ into which the elements of the complex $\boldsymbol{W}$, corresponding to a relation between the elements b, are divided, when the 1-cells corresponding to elements of b are taken away.

The boundary of this component $\boldsymbol{C}_i$ will consist of a certain number of closed curves representing relations between the elements b. The relation represented by the exterior boundary is called $T(b) = 1$. The others are called $T_l(b) = 1$, $l = 1, \cdots, m$. The point P where we start reading the product $T(b)$ can be joined on the component plus its boundary to the corresponding point for $T_1(b)$ by an arc representing the product $\phi_1(a_i, b)$; next P can be joined by an arc, not crossing the first arc and representing the product $\phi_2(a_i, b)$ to the corresponding point for $T_2(b)$, and so on. We finally find a relation (5), consequence of one set of relations $R_i(a_i, b) = 1$ and such that the relations $T_l(b) = 1$, are a consequence of the existence of complexes in the interior of the closed curves representing these relations. The original relation is a consequence of this relation (5) just found and of one or more other relations between the elements b that can be represented by complexes containing a fewer number of 2-cells than $\boldsymbol{W}$. It follows that Theorem 2 can be proved by an induction on the number of 2-cells in $\boldsymbol{W}$.

---

* The theorem proved by Schreier (*loc. cit.*), is practically identical with Corollary 1 together with Theorem 1 in those cases where Corollary 1 can be applied.

THEOREM 3. *If we divide the sets of generators* a_i *each into two parts,* a_{1i} *and* a_{2i}, *then any relation* $W(a_{1i}, b) = 1$ *in* $\mathfrak{G}$ *between the elements* b *and* a_{1i}, $i = 1, 2, \cdots$, *is a consequence of the set of relations, that result from the elimination of the elements* a_{2i} *from the relations* $R(a_i, b) = 1$ *and all relations in* $\mathfrak{G}$ *between the elements* b.

Proof. As in the proof of Theorem 2, we determine all components $\boldsymbol{C}_i$ of the complex $\boldsymbol{W}$ representing the relation $W = 1$. Any component $\boldsymbol{C}_l$, having at least one 1-cell in common with the boundary of $\boldsymbol{W}$, represents a relation between the elements of the set b and those of a set a_{1i}. This relation, according to Lemma 2, is a consequence of the relations $R_i(a_i, b) = 1$ and of relations represented by components $\boldsymbol{C}_i$ of $\boldsymbol{C}$ of which the boundary does not have a 1-cell in common with the boundary of $\boldsymbol{W}$. These relations are relations between the elements b alone. But the original relation follows from the relations expressed by the boundaries of components $\boldsymbol{C}_l$ of the type just considered and of some more relations represented by components having no 1-cell in common with the boundary of $\boldsymbol{W}$, that means of some more relations between the elements b alone. Thus Theorem 3 is proved.

COROLLARY 2. *Any relation* $W(a_1, b) = 1$ *valid in* $\mathfrak{G}$ *between the elements* a_1 *and* b *is a consequence of the relations* $R_1(a_1, b) = 1$ *and all relations between the elements* b *valid in* $\mathfrak{G}$.

THE JOHNS HOPKINS UNIVERSITY.

THE INTEGERS REPRESENTED BY SETS OF TERNARY QUADRATIC FORMS.*

By A. Adrian Albert.

1. *Introduction.* One of the most interesting topics in the theory of numbers is the study of the question of what integers are represented by positive ternary quadratic forms. Few general theorems are known in this subject. In fact, as L. E. Dickson has indicated, most of the forms are irregular.†

In the present paper a consideration is made of a different type of problem ‡ yet one that throws a good deal of light on the above topic. The problem of determining all the integers represented by the *set* $\Sigma(d)$ of *all positive ternary quadratic forms* of the *same determinant* d is studied here and a complete solution is obtained. Moreover the results have the following remarkably simple form.

We may write any two positive integers in their unique form

$$d = \gamma^2\delta, \quad a = \rho^2\sigma, \tag{1}$$

where δ and σ have no square factors. Then it is shown here that $\Sigma(d)$ represents every integer a not of the form

$$a = \rho^2\sigma, \quad \sigma = \alpha\delta, \quad \alpha = 8n + 7, \tag{2}$$

such that α is prime to d and is such that the Jacobi symbol

$$(p \mid \alpha) = +1 \tag{3}$$

* Presented to the Society February 28, 1931. Received by the Editors in July, 1932.

† For the definition of regularity see our Section 7. For L. E. Dickson's quoted paper see the *Annals of Mathematics*, Vol. 28 (1926-27), pp. 333-341.

‡ Note added June 30, 1932. Due to my recent activity in the study of linear algebras I have been unable to prepare the present paper (of which an abstract giving explicit results appeared in 1931 in the *Bulletin of the American Mathematical Society*) for publication until now. Since it was written B. W. Jones has proved (in the *Transactions of the American Mathematical Society*, Vol. 33 (1931), pp. 92-124) that every genus of positive ternaries is regular. The problem solved by Jones is a different one although quite close to the one here considered. Moreover it is a much more complicated problem (since a genus of ternaries is itself a complicated notion) so that the results of Jones are not as simple as those given here. As my theory was obtained *independently* of the theory of Jones, and as our two problems are really distinct, I believe my paper to be still of the same interest as before the publication of the papers by B. W. Jones.

for every prime factor p of d. This result is a real generalization of the case $d = 1$ in which case $\Sigma(1)$, consisting of a single form in the sense of equivalence, is well known to represent every a not of the form $4^k(8n+7)$, that is all integers a not of the form (1) with $\delta = d = 1$.

The above simple result on sets $\Sigma(d)$ is easily shown to imply that every $\Sigma(d)$ is regular in the Dickson sense. Moreover it is shown that every $\Sigma(d)$ represents in particular no integer $\delta(8nd - 1)$ so that no $\Sigma(d)$ represents all positive integers. However if $\Sigma(n, d)$ is the set of all positive n-aries of the same determinant d, $n \geqq 4$, then every $\Sigma(n, d)$ represents all positive integers.*

2. *Preliminary theory.* We shall consider ternary quadratic forms

$$(4) \qquad \phi \equiv \phi(x, y, z) = ax^2 + by^2 + cz^2 + 2ryz + 2sxz + 2txy,$$

where a, b, c, r, s, t are integers and x, y, z integer variables. A form ϕ is called positive if $\phi \geqq 0$ for all integers x, y, z and if $\phi = 0$ if and only if $x = y = z = 0$. It is well known that ϕ is positive if and only if

$$(5) \qquad a > 0, \quad ab - t^2 > 0, \quad d > 0,$$

where the determinant d of ϕ has the value

$$(6) \qquad d = c(ab - t^2) + 2rst - ar^2 - bs^2.$$

An integer q is said to be represented by ϕ if there exist integers α, β, γ for which $\phi(\alpha, \beta, \gamma) = q$. If the greatest common divisor of α, β, γ is unity then the representation is proper. Moreover when q is represented properly by ϕ there exists † a transformation of determinant unity replacing ϕ by an equivalent form

$$(7) \qquad \Phi(X, Y, Z) \equiv AX^2 + BY^2 + CZ^2 + 2RYZ + 2SXZ + 2TXY,$$

---

* Note added January 19, 1933. This paper in its present form is a revision in accordance with the suggestions of the referee concerning the paper originally offered for publication to the Editors of this Journal. It is approximately three-eighths shorter than the original because of two major changes. First, certain suggestions of the referee caused me, by implication, to make a reduction to the case d an odd prime with a resulting great saving in space. Secondly, this reduction enabled me to use the referee's elegant short proof of the necessity part of Theorem 9 of this paper instead of my original much longer proof of the same theorem but for the case of a general d. I thank him for the opportunity to use his shorter proof.

† For this and other properties of ternary quadratic forms given in this section one may see L. E. Dickson's *Studies in the Theory of Numbers*, Chicago, 1930.

with $A = q$. Hence when a positive * integer a is represented by some form in $\Sigma(d)$ there must exist integers b, c, r, s, t for which the corresponding form (4) has determinant d satisfying (6) and hence

$$d + ar^2 + bs^2 - 2rst = (ab - t^2)c, \tag{8}$$

$$d + ar^2 + bs^2 - 2rst \equiv 0 \pmod{c}, \tag{9}$$

with $ab - t^2 > 0$. Conversely if $a > 0$, $d > 0$ and there exists a set of integers b, c, r, s, t for which (8) is satisfied with $ab - t^2$ positive then the corresponding form (4) has determinant d, is in $\Sigma(d)$, and represents a properly for $x = 1$, $y = z = 0$.

An integer a represented properly by some form (4) of determinant d appears as the coefficient of x^2 in some form of the set $\Sigma(d)$ and conversely. This justifies the following definition †

DEFINITION. *An integer a is called a coefficient of $\Sigma(d)$ if a is represented properly by a form of $\Sigma(d)$.*

We have then proved that a is a coefficient of $\Sigma(d)$ if and only if there exist integers b, r, s, t, c satisfying (8) and with $ab - t^2 > 0$. But the condition (9) is equivalent to (8) since evidently (8) implies (9) while if (9) is satisfied we may define c as the quotient of the left member of (9) by $ab - t^2$ and (8) is satisfied. We therefore have the criterion

THEOREM 1. *An integer $a > 0$ is a coefficient of a set $\Sigma(d)$ if and only if there exist integers b, c, r, s, t for which $ab - t^2$ is positive and one of the equivalent conditions* (8) *and* (9) *is satisfied.*

3. *Some general results.* We shall first obtain a result of great importance for the case δ even. Let a be a coefficient of $\Sigma(d)$ so that (8) is satisfied. If one of b and c is even then, by interchanging y and z, b and c if necessary, we may take c even. If both are odd the transformation

$$x = X, \quad y = Y + Z, \quad z = Z \tag{10}$$

of determinant unity replaces (4) by an equivalent form (7) in which

$$A = a, \quad B = b, \quad C = b + c + 2r, \quad R = r + b, \quad S = s + t, \quad T = t, \tag{11}$$

* In all our subsequent work $a > 0$, $d > 0$ without mentioning this fact. It will also be unnecessary to write $ab - t^2 > 0$ as if a is represented properly by $\Sigma(d)$ then $ab - t^2 > 0$ while conversely all $ab - t^2$ used will be obtained as the products or quotients of *given positive integers*.

† We make this definition to avoid the constant repetition of the phrase *represented properly*.

and with $C = b + c + 2r$ even. Hence we may always take c even, $c = 2c_1$, whence (8) becomes

$$(12) \qquad 2d + 2ar^2 + 2bs^2 - 2rs(2t) = [2a \cdot 2b - (2t)^2]c_1.$$

Hence $2a$ is a coefficient of $\Sigma(2d)$.

Conversely let $2a$ be a coefficient of $\Sigma(2d)$. Then there exist integers b, c, r, s, t, for which (4) has determinant $2d$ with

$$(13) \qquad 2d + 2ar^2 + bs^2 - 2rst = (2a \cdot b - t^2)c.$$

If one of s and t is even we may take s even by the argument above. If both are odd then the use of (10), (11) with $S = s + t$ gives S even. Hence we may always take s even, $s = 2s_0$.

If c is even, $c = 2c_0$ then (13) becomes

$$(14) \qquad 2d + 2ar^2 + 4bs_0^2 - 4rs_0t = 2(2ab - t^2)c_0,$$

so that

$$(15) \qquad d + ar^2 + 2b \cdot s_0^2 - 2rs_0t = (a \cdot 2b - t^2)c_0,$$

and, by Theorem 1, a is a coefficient of $\Sigma(d)$.

Let then c be odd. Since s is even (13) implies that $2ab - t^2$ is even. But then t is even. Since both s and t are now even the above argument and the use of (10) if necessary imply that we may take C even and yet S and T even. Hence we may always take s even, c even so that (14) holds and a is a coefficient of $\Sigma(d)$.

THEOREM 2. *An integer $2a$ is a coefficient of $\Sigma(2d)$ if and only if a is a coefficient of $\Sigma(d)$.*

If k is any integer and a is a coefficient of $\Sigma(d)$ equation (8) implies that

$$(16) \qquad k^2d + a(kr)^2 + b(ks)^2 - 2(kr)(ks)t = (ab - t^2)k^2c,$$

and, by Theorem 1, we have

THEOREM 3. *If a is a coefficient of $\Sigma(d)$ then a is a coefficient of $\Sigma(k^2d)$ for every integer k.*

But (16) may also be written

$$(17) \qquad k^2d + k^2a \cdot r^2 + b(ks)^2 - 2r(ks)(kt) = (k^2a \cdot b - k^2t^2)c,$$

so that Theorem 1 gives

LEMMA 1. *If a is a coefficient of $\Sigma(d)$ then k^2a is a coefficient of $\Sigma(k^2d)$ for every integer k.*

We shall also prove a less obvious theorem, the converse of Lemma 1. Let p be a prime, n a positive integer. If in (4) the g. c. d. of s and p^n is p^m, the g. c. d. of t and p^n is p^μ then, by interchanging y and z if necessary (and hence s and t), we may always take $\mu \leq m$. In the linear congruence

$$t\xi + s \equiv 0 \pmod{p^n} \tag{18}$$

the g. c. d. p^μ of p^n and t divides s. Hence there exists an integer ξ for which (18) is satisfied. For such an integer ξ the transformation

$$x = X, \quad y = Y + \xi Z, \quad z = Z \tag{19}$$

of determinant

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & \xi \\ 0 & 0 & 1 \end{vmatrix} = 1$$

replaces (4) by an equivalent form (7) in which

$$A = a,\ B = b,\ C = b\xi^2 + 2r\xi + c,\ R = b\xi + r,\ S = t\xi + s,\ T = t, \tag{20}$$

and $A = a$, $S = t\xi + s \equiv 0 \pmod{p^n}$.

LEMMA 2. *Let p be a prime, n a positive integer, a be a coefficient of $\Sigma(d)$. Then* (8) *is satisfied with $s \equiv 0 \pmod{p^n}$.*

As an immediate corollary we have

LEMMA 3. *Let p, n, a, d be as in Lemma 2. Then* (8) *is satisfied with $t \equiv 0 \pmod{p^n}$.*

We may now prove

LEMMA 4. *Let p be a prime and p^2a be a coefficient of $\Sigma(p^2d)$. Then a is a coefficient of $\Sigma(d)$.*

For by Theorem 1 there exist integers b, c, r, s, t for which

$$p^2d + p^2ar^2 + bs^2 - 2rst = (p^2a \cdot b - t^2)c, \tag{21}$$

and, by Lemma 2, with $s \equiv 0 \pmod{p^2}$. But then (21) implies that $c(p^2ab - t^2) \equiv 0 \pmod{p^2}$. If $t \not\equiv 0 \pmod{p}$ then $c \equiv 0 \pmod{p^2}$, $c = c_0p^2$, $s = s_0p^2$, and if we define $b_0 = p^2b$, (21) implies

$$p^2d + p^2ar^2 + p^2b_0s_0^2 - 2rs_0tp^2 = (ab_0 - t^2)c_0p^2,$$

so that

(22) $$d + ar^2 + b_0 s_0^2 - 2rs_0 t = (ab_0 - t^2)c_0,$$

and a is a coefficient of $\Sigma(d)$.

Let then $t \equiv 0 \pmod{p}$, $t = t_1 p$, $s = s_1 p$. Then (21) gives

$$p^2 d + p^2 ar^2 + p^2 bs_1^2 - 2rs_1 t_1 p^2 = (ab - t_1^2)p^2 c,$$

whence

(23) $$d + ar^2 + bs_1^2 - 2rs_1 t_1 = (ab - t_1^2)c,$$

and again a is a coefficient of $\Sigma(d)$ so that Lemma 4 is proved.

Let $k^2 a$ be a coefficient of $\Sigma(k^2 d)$. If $k = k_0 p$ where p is a prime then Lemma 4 implies that $k_0^2 a$ is a coefficient of $\Sigma(k_0^2 d)$. A repetition of this process combined with Lemma 1 evidently gives

THEOREM 4. *An integer $k^2 a$ is a coefficient of $\Sigma(k^2 d)$ if and only if a is a coefficient of $\Sigma(d)$.*

We may write any integer a in the form $a = \rho^2 \sigma$ where σ has no square factor. An almost obvious necessary condition that a be a coefficient of $\Sigma(d)$ in view of our theorems is

THEOREM 5. *An integer $a = \rho^2\sigma$, σ with no square factor, is a coefficient of $\Sigma(d)$ only when σ is a coefficient of $\Sigma(d)$.*

For by Theorem 3 if a is a coefficient of $\Sigma(d)$ then $a = \rho^2\sigma$ is a coefficient of $\Sigma(\rho^2 d)$. By Theorem 4 σ is a coefficient of $\Sigma(d)$.

Let then $a = \rho^2\sigma$ be a positive integer represented by $\Sigma(d)$. Then there is a form (4) of determinant d and integers α, β, γ such that $\phi(\alpha, \beta, \gamma) = a$. Let the g. c. d. of α, β, γ be ξ so that $\alpha = \xi\alpha_1$, $\beta = \xi\beta_1$, $\gamma = \xi\gamma_1$ where $\alpha_1, \gamma_1, \beta_1$ are relatively prime. Then $a = \xi^2\phi(\alpha_1, \beta_1, \gamma_1)$. Write $\phi(\alpha_1, \beta_1, \gamma_1) = \eta^2\epsilon$ where ϵ has no square factor. Then $(\xi\eta)^2\epsilon = \rho^2\sigma$ whence $\epsilon = \sigma$ and $\eta^2\epsilon = \eta^2\sigma = \phi(\alpha_1, \beta_1, \gamma_1)$ so that $\eta^2\sigma$ is a coefficient of $\Sigma(d)$. By Theorem 5 σ is a coefficient of $\Sigma(d)$.

Conversely let σ be a coefficient of $\Sigma(d)$. Then there is a form $\phi(x, y, z)$ of determinant d and with σ as the coefficient of x^2. Obviously $a = \rho^2\sigma = \phi(\rho, 0, 0)$ is represented by $\Sigma(d)$. Our problem of determining all integers a represented by sets $\Sigma(d)$ has therefore been simplified by

THEOREM 6. *An integer $a = \rho^2\sigma$, σ with no square factor, is represented by a set $\Sigma(d)$ if and only if σ is a coefficient of $\Sigma(d)$.*

We shall next obtain one more general result of great importance in our work. We let p be a prime divisor of d which does not divide a. We may then prove

LEMMA 5. *Let a be a coefficient of $\Sigma(d)$. Then* (8) *is satisfied with*

$$c \equiv s \equiv t \equiv 0 \pmod{p}. \tag{24}$$

For by Lemma 3 we may take $t \equiv 0 \pmod{p}$. Let ξ be chosen so that $\xi a + s \equiv 0 \pmod{p}$. This may be accomplished since a is not divisible by the prime p. The transformation

$$x = X + \xi Y, \quad y = Y, \quad z = Z \tag{25}$$

replaces (4) by an equivalent form (7) in which

$$A = a, \quad B = b, \quad C = a\xi^2 + 2s\xi + c, \quad R = r + t\xi, \quad S = a\xi + s, \quad T = t,$$

so that S is divisible by p, $T = t$. Hence we may take $s \equiv t \equiv 0 \pmod{p}$. If one of b and c is divisible by p then we may interchange them if necessary and (24) is satisfied. If $b \not\equiv 0 \pmod{p}$ then there exists an integer η satisfying $b\eta + r \equiv 0 \pmod{p}$ and the transformation (12) replaces (4) by an equivalent form (7) in which

$$A = a,\ B = b,\ C = b\eta^2 + 2r\eta + c,\ R = b\eta + r,\ S = t\eta + s,\ T = t \tag{26}$$

is satisfied so that $S \equiv T \equiv 0 \pmod{p}$. But $bC = (b\eta + r)^2 + (bc - r^2)$, $b\eta + r \equiv 0 \pmod{p}$. Also $d + bs^2 + ct^2 - 2rst = (bc - r^2)a \equiv 0 \pmod{p}$ since $d \equiv s \equiv t \equiv 0 \pmod{p}$. Hence also $bc - r^2 \equiv 0 \pmod{p}$ so that $bC \equiv 0 \pmod{p}$. It follows that $C \equiv 0 \pmod{p}$ and the Lemma is proved.

Hence we have shown that if p is a prime divisor of d but not of a coefficient a of $\Sigma(d)$ then (8) is satisfied with $c = c_0 p$, c_0 an integer. Then

$$p(d + ar^2 + bs^2 - 2rst) = c_0(p^2ab - p^2t^2). \tag{27}$$

Let us define $t_1 = pt$, $b_1 = pb$. Then (27) is equivalent to

$$pd + pa \cdot r^2 + b_1 s^2 - 2rst_1 = c_0[(pa)(b_1) - t_1^2], \tag{28}$$

and we have proved

THEOREM 7. *Let a prime p divide d and not a. Then if a is a coefficient of $\Sigma(d)$ the integer pa is a coefficient of $\Sigma(pd)$.*

4. *Reduction to the case $d = \delta$.* Let $d = \gamma^2\delta$ be odd and δ have no square factor. In the present section we shall prove a theorem which will later be seen to have reduced our problem to essentially the case $d = \delta$, δ odd.

We suppose first that ν is a positive integer with no square factor and let a be a coefficient of $\Sigma(\nu^2 d)$. By Theorem 1 we may write

$$\nu^2 d + ar^2 + bs^2 - 2rst \equiv 0 \pmod{ab - t^2}. \tag{29}$$

Let the g. c. d. of ν and $ab - t^2$ be ν_0 so that

$$(30) \qquad ab - t^2 \equiv 0 \pmod{\nu_0}, \qquad \nu = \nu_0\nu_1,$$

where ν_1 is prime to $ab - t^2$ since ν has no square factor. Then there exists an integer g for which $g\nu_1 \equiv 1 \pmod{ab - t^2}$. Hence if $s_1 = gs$, $r_1 = gr$ then (29) implies

$$(31) \qquad \nu_0^2 d + ar_1^2 + bs_1^2 - 2r_1s_1t \equiv 0 \pmod{ab - t^2}.$$

By Theorem 1 a is a coefficient of $\Sigma(\nu_0^2 d)$ and in fact there exists an integer c_1 defined by (31) and satisfying

$$(32) \qquad \nu_0^2 d + ar_1^2 + bs_1^2 - 2r_1s_1t = (ab - t^2)c_1,$$

from which we may write

$$(33) \qquad \nu_0^2 d + ar_1^2 + c_1t^2 - 2r_1ts_1 = (ac_1 - s_1^2)b.$$

Let the g. c. d. of ν_0 and $ac_1 - s_1^2$ be ν_2 so that

$$(34) \qquad ac_1 - s_1^2 \equiv 0 \pmod{\nu_2}, \qquad \nu_0 = \nu_2\nu_3,$$

and ν_3 is prime to $ac_1 - s_1^2$. As above (33) implies that if $g_0\nu_3 \equiv 1$ (mod $ac_1 - s_1^2$), $r_2 = g_0r_1$, $t_2 = tg_0$ then

$$(35) \qquad \nu_2^2 d + ar_2^2 + c_1t_2^2 - 2r_2t_2s_1 \equiv 0 \pmod{ac_1 - s_1^2},$$

and a is a coefficient of $\Sigma(\nu_2^2 d)$. Now (34) is satisfied but not necessarily (30) with b replaced by b_1 defined as the quotient of the left member of the congruence (35) by its modulus. However we have thus determined a decreasing sequence of positive integers $\nu, \nu_0, \nu_2, \cdots$ which terminates when at some stage we obtain an integer N for which

$$(36) \qquad N^2 d + ar^2 + bs^2 - 2rst = (ab - t^2),$$

with

$$(37) \qquad \nu \equiv ab - t^2 \equiv ac - s^2 \equiv 0 \pmod{N}.$$

Since a is prime to ν and hence to N there exists an integer ξ such that $a\xi + t \equiv 0 \pmod{N}$. The transformation

$$(38) \qquad x = X + \xi Y, \quad y = Y, \quad z = Z$$

replaces (4) by an equivalent form (7) in which

$$(39) \quad A = a,\ B = a\xi^2 + 2t\xi + b,\ C = c,\ R = r + \xi s,\ S = s,\ T = \xi a + t,$$

while (4) has determinant $N^2 d$ and satisfies (37). But then $T \equiv 0 \pmod{N}$,

$aC - S^2 = ac - s^2 \equiv 0 \pmod{N}$. Also $aB = (a\xi + t)^2 + (ab - t^2) \equiv 0 \pmod{N}$. Evidently then $aB - T^2 \equiv B \equiv 0 \pmod{N}$ since a is prime to N.

Hence we have shown that we may assume (37) as well as

$$b \equiv t \equiv 0 \pmod{N}. \tag{40}$$

Similarly there exists an integer η for which $\eta a + s \equiv 0 \pmod{N}$ so that the transformation

$$x = X + \eta Z, \quad y = Y, \quad z = Z \tag{41}$$

replaces (4) by a form (7) in which

$$A = a,\ B = b,\ C = a\eta^2 + 2s\eta + c,\ R = r + \eta t,\ S = \eta a + s,\ T = t, \tag{42}$$

so that $S \equiv T \equiv B \equiv 0 \pmod{N}$ while $aC = (a\eta + s)^2 + (ac - s^2) \equiv 0 \pmod{N}$ and hence $C \equiv 0 \pmod{N}$. From (36) in capitals we have also $aR^2 \equiv 0 \pmod{N}$ whence $R \equiv 0 \pmod{p}$ and we have proved

LEMMA 5. *Let a be prime to ν, an integer with no square factor, and let a be a coefficient of $\Sigma(\nu^2 d)$. Then there is a factor N of ν such that a is a coefficient of $\Sigma(N^2 d)$ with a corresponding equation* (8) *with d replaced by $N^2 d$ and*

$$b \equiv c \equiv r \equiv s \equiv t \equiv 0 \pmod{N}. \tag{43}$$

But now we have

$$b = b_1 N, \quad c = c_1 N, \quad r = r_1 N, \quad s = s_1 N, \quad t = t_1 N, \tag{44}$$

and

$$N^2 d + N^2 a r_1^2 + N^3 b_1 s_1^2 - 2 r_1 s_1 t_1 N^3 = (ab - t^2) N c_1, \tag{45}$$

so that

$$(Nd) + (Na) r_1^2 + b_1 s^2 - 2 r_1 s t = (Na \cdot b_1 - t^2) c_1 \tag{46}$$

and Na is a coefficient of $\Sigma(Nd)$. We have proved

LEMMA 6. *Let ν have no square factor and a be prime to ν and a coefficient of $\Sigma(\nu^2 d)$. Then there exists a factor N of ν such that Na is a coefficient of $\Sigma(Nd)$.*

We shall assume now that ν is a prime p so that if a is a coefficient of $\Sigma(p^2 d)$ then either $N = 1$ or p and either a is a coefficient of $\Sigma(d)$ or pa is a coefficient of $\Sigma(pd)$. Assume moreover that d is divisible by p. Then in the latter case we use (46) and obtain

$$pd + par^2 + bs^2 - 2rst = (pab - t^2)c, \tag{47}$$

with $s \equiv t \equiv 0 \pmod{p}$. Evidently $pd + bs^2 - 2rst + t^2 c \equiv 0 \pmod{p^2}$, so that $ap(r^2 - bc) \equiv 0 \pmod{p}$ and hence $r^2 - bc \equiv 0 \pmod{p}$.

If $c \equiv 0 \pmod{p}$ we write $s = s_1 p$, $c = c_1 p$ and obtain

$$d + ar^2 + (pb)s_1{}^2 - 2rs_1 t = (a \cdot pb - t^2)c.$$

Then a is a coefficient of $\Sigma(d)$.

If $c \not\equiv 0 \pmod{p}$ we choose η so that $c\eta + r \equiv 0 \pmod{p}$ and use the transformation

$$x = X, \quad y = Y, \quad z = \eta Y + Z \tag{48}$$

to obtain a new form (7) of $\Sigma(pd)$ in which

$$A = a,\ B = c\eta^2 + 2r\eta + b,\ C = c,\ R = \eta c + r,\ S = s,\ T = s\eta + t, \tag{49}$$

so that $S \equiv T \equiv 0 \pmod{p}$ as for s and t and $cB = (c\eta + r)^2 + (bc - r^2)$ is also divisible by p. But then $B \equiv 0 \pmod{p}$. The interchange of B and C simultaneously with that of S and T gives a form ϕ of the set $\Sigma(pd)$ in which $s \equiv t \equiv c \equiv 0 \pmod{p}$ so that, by the above argument, a is a coefficient of $\Sigma(d)$. We have proved

Lemma 7. *Let a prime divisor p of d be not a divisor of a. Then a is a coefficient of $\Sigma(p^2 d)$ only if a is a coefficient of $\Sigma(d)$.*

We next suppose that $a = \sigma$ has no square factor and $a = a_1 p$, $d = p^3 d_1$. Then if p is a prime

$$d_1 p^3 + ar^2 + bs^2 - 2rst = (ab - t^2)c, \tag{50}$$

so that if $b_1 = pb$ then

$$p^4 d_1 + a_1 (pr)^2 + b_1 s^2 - 2(pr)st = (ab_1 - t^2), \tag{51}$$

and a_1 is a coefficient of $\Sigma(p^4 d_1)$. By Lemma 7 a_1 is a coefficient of $\Sigma(p^2 d_1)$.

Suppose now that $d = \gamma^2 \delta$, $a = \sigma = \alpha\delta$ where σ has no square factor, α is prime to d. Then if γ has a factor p in common with δ we have $d = p^3 d_1$, $a = a_1 p = \alpha\delta_1 p$, $\delta = \delta_1\ p$ and, as above, $\alpha\delta_1$ is a coefficient of $\Sigma(p^2 d_1)$ where $p^2 d_1 = \gamma^2 \delta_1$. But then $a_1 = \alpha\delta_1$ is a coefficient of $\Sigma(\gamma^2\delta_1)$, that is we have replaced δ by a factor δ_1 of δ. Also δ_1 is prime to p.

Hence we may take δ_0 prime to γ, $a_0 = \alpha\delta_0$ to be a coefficient of $\Sigma(\gamma^2\delta_0)$ without loss of generality. Let next γ have a square factor p^2, p a prime, $\gamma = \gamma_0 p$ where $d_0 = \gamma_0{}^2\delta_0$ is divisible by p. Then by Lemma 7 a_0 is a coefficient of $\Sigma(\gamma_0{}^2\delta_0)$. Hence we may take $\gamma = \nu$ and have ν an integer with no square factor prime to δ_0 and hence to $\alpha\delta_0$.

Now $\alpha\delta_0$ is a coefficient of $\Sigma(\nu^2\delta_0)$. Since α is prime to $\nu^2\delta_0$ and ν is prime to δ_0 the integer $\nu\alpha\delta_0$ has no square factor. By Lemma 6 there exists a factor N of ν such that $N\alpha\delta_0$ is a coefficient of $\Sigma(N\delta_0)$. Let $\Delta = N\delta_0$.

Then $\alpha\Delta$ has no square factor and is a coefficient of $\Sigma(\Delta)$. This completes the proof of

THEOREM 8. *Let $\sigma = \alpha\delta$ have no square factor, α be prime to d, and σ be a coefficient of $\Sigma(d)$, where $d = \gamma^2\delta$. Then there is a factor Δ of d such that $\alpha\Delta$ has no square factor and is a coefficient of $\Sigma(\Delta)$.*

5. *The case d odd, $d = \delta$. We shall let d be an odd integer with no square factor in this section.*

Let b be chosen so that $ab - 1$ is an odd prime not a divisor of ad and such that the Legendre symbol

$$(-ad \mid p) = +1. \tag{52}$$

Then $-ad$ is a quadratic residue of p and there exists an integer s for which

$$-ad \equiv s^2 \pmod{p}. \tag{53}$$

But $ab \equiv 1 \pmod{p}$ and p is prime to a. Hence $-ad \equiv abs^2 \pmod{p}$, $-d \equiv bs^2 \pmod{p}$, and

$$d + bs^2 \equiv 0 \pmod{ab - 1}, \quad ab - 1 = p > 0. \tag{54}$$

By Theorem 1 with $r = 0$, $t = 1$ we have

LEMMA 8. *An integer a is a coefficient of $\Sigma(d)$ if there exists an odd prime $p = ab - 1$ not a factor of d and such that $(-ad \mid p) = 1$.*

We shall repeatedly use *Dirichlet's theorem on the primes in an arithmetic progression.* The progression

$$(8a_1d)m + (4a_1d - 1) \tag{55}$$

evidently has relatively prime coefficients and hence, by Dirichlet's theorem, contains a prime p for n properly chosen. Let

$$a = 2a_1, \quad b = 2b_1, \quad b_1 = d(2n + 1), \tag{56}$$

so that

$$ab - 1 = 4a_1d(2n + 1) - 1 = (8a_1d)n + (4a_1d - 1) = p. \tag{57}$$

Evidently p is prime to ad. Also

$$(-ad \mid p) = (-2 \mid p)(a_1d \mid p) = (-1)^\beta, \tag{58}$$

where

$$\begin{cases} \beta = \dfrac{p-1}{2} + \dfrac{p^2-1}{8} + \dfrac{a_1d-1}{2}\,\dfrac{p-1}{2} + \dfrac{a_1d-1}{2} \\[2ex] \quad = \dfrac{p-1}{2}\left(\dfrac{p+5}{4}\right) + \dfrac{p+1}{2}\,\dfrac{a_1d-1}{2}. \end{cases} \tag{59}$$

But $p = 4m - 1$, $m = (2n+1)a_1 d$ is odd and hence $p + 1 \equiv 0 \pmod 4$, $p + 5 = 4(m+1) \equiv 0 \pmod 8$. Hence β is a sum of even integers and is even. By Lemma 8 we have

LEMMA 9. *Every even a is a coefficient of $\Sigma(d)$.*

where we are assuming that $a = \sigma$ has no square factor.

Next let $\sigma = a$ be odd and not divisible by d. Then there is an odd prime factor q of d not dividing a so that $d = Dq$. Suppose first that $aD + 1 \equiv 0 \pmod 4$. Let η be a quadratic non-residue of q. The congruences

$$(60) \qquad p \equiv -1 \pmod{aD}, \qquad p \equiv \eta \pmod q,$$

have relatively prime moduli and hence, by the *Chinese remainder theorem,* have a common solution p_0 such that every solution has the form

$$(61) \qquad p = aqDn + p_0 = adn + p_0.$$

Evidently ad is prime to p_0. Hence n may be chosen so that p is a prime. By (60) $p = aDm - 1 = ab - 1$ if $b = Dm$. Also

$$(62) \qquad \begin{aligned}(-ad \mid p) &= (-1)^{(p-1)/2 + [(p-1)/2]\,[(ad-1)/2]}(p \mid ad) \\ &= (p \mid q)(-1)^{[(p-1)/2]\,[(ad+1)/2] + (aD-1)/2}.\end{aligned}$$

But $p \equiv \eta \pmod q$, $(\eta \mid q) = -1$ and $aD + 1 \equiv 0 \pmod 4$ so that $aD - 1 \equiv 2 \pmod 4$. Hence

$$(63) \qquad (-ad \mid p) = (-1)(-1)^{(aD-1)/2} = +1.$$

By Lemma 8 a is a coefficient of $\Sigma(d)$.

Next let $aD + 1 \equiv 2 \pmod 4$. Then the congruences

$$p \equiv -1 \pmod{aD}, \qquad p \equiv 1 \pmod{4q}$$

have relatively prime moduli and a common solution

$$(64) \qquad p = 4adn + p_0 = aDm - 1 = ab - 1,$$

if $b = Dm$. We may evidently choose p to be a prime and, as before have

$$(65) \qquad (-ad \mid p) = (-1)^{[(p-1)/2]\,[(ad+1)/2] + (aD-1)/2}(p \mid q).$$

But $p \equiv 1 \pmod q$, $p \equiv 1 \pmod 4$, $aD - 1 \equiv 0 \pmod 4$. Hence $(-ad \mid p) = 1$. We have proved, by the use of Lemma 9 if a is even,

LEMMA 10. *Let a be not divisible by d. Then a is a coefficient of $\Sigma(d)$.*

Suppose now that a is divisible by d, $a = \alpha d$ but $\alpha \not\equiv 7 \pmod 8$. If $\alpha \equiv 1 \pmod 4$ then we may take $p = ab - 1$ to be a prime

$$(66) \qquad p = ad(4m+2) - 1 = \alpha d^2(4m+2) - 1 \equiv 1 \pmod{4},$$

by proper choice of m. But in this case

$$(67) \qquad \begin{aligned}(-ad \mid p) &= (-1)^{(p-1)/2} (\alpha \mid p) \\ &= (-1)^{[(p-1)/2]\,[(\alpha+1)/2]+(\alpha-1)/2} = (-1)^{[(p+1)/2]\,[(\alpha+1)/2]-1}.\end{aligned}$$

But $(p+1)(\alpha+1) = 4k$ where k is odd so that $(-ad \mid p) = 1$ as desired. Then a is a coefficient of $\Sigma(d)$.

Let next $\alpha \equiv 3 \pmod{4}$ so that $\alpha \equiv 3 \pmod{8}$ if $\alpha \not\equiv 7 \pmod{8}$. In this case $\alpha = 8m + 3$

$$(68) \qquad (2 \mid \alpha) = (-1)^{(\alpha^2-1)/8} = -1,$$

since $\alpha^2 - 1 \equiv 9 - 1 \equiv 8 \pmod{16}$. But now we take $p = adn - 2$ chosen to be a prime. Then we write

$$(69) \qquad 2p = 2adn - 4 = (2dn)a - 2^2 = ab - t^2,$$

where $t = 2$, $b = 2dn$. Now

$$(70) \qquad \begin{aligned}(-ad \mid p) &= (-\alpha \mid p) = (-1)^{[(p-1)/2]\,[(\alpha+1)/2]} (p \mid \alpha) \\ &= (-1)^{[(p-1)/2]\,[(\alpha+1)/2]} (2 \mid \alpha)(2p \mid \alpha) = (-1)^{[(p-1)/2]\,[(\alpha+1)/2]+1} (-1 \mid \alpha) \\ &= (-1)^{[(p-1)/2]\,[(\alpha+1)/2]+1+(\alpha-1)/2} = (-1)^{[(p+1)/2]\,[(\alpha+1)/2]} = 1,\end{aligned}$$

since $\alpha + 1 \equiv 0 \pmod{4}$. Hence $-ad$ is a quadratic residue of p,

$$(71) \qquad -ad \equiv \eta^2 \pmod{p}.$$

Evidently one of p and $\eta - p$ is odd and we may take η so that $-ad \equiv \eta^2 \pmod{2p}$. But then $-ad \equiv \eta^2 \pmod{ab - 4}$. The integer a is prime to $ab - 4$ since a is odd and there exists an r such that $ar \equiv \eta \pmod{ab-4}$. Then $-ad \equiv a^2r^2 \pmod{2p}$ and hence $d + ar^2 \equiv 0 \pmod{ab - t^2}$, $t = 2$. By Theorem 1 a is a coefficient of $\Sigma(d)$.

Suppose finally that $\alpha \equiv 7 \pmod{8}$ but let there exist an odd prime factor q of d such that

$$(72) \qquad (q \mid \alpha) = -1.$$

Let $d = qd_1$ and write

$$(73) \qquad p = an + (\alpha d_1 - q).$$

Every factor of a divides one of the relatively prime integers α, d_1, q and hence is prime to $\alpha d_1 - q$. Hence n may be selected so that p is a prime. Then if

$$(74) \qquad t = q, \quad b = qn + 1 = tn + 1,$$

we have

$$(75) \qquad \begin{aligned}ab - t^2 &= a(qn+1) - q^2 = \alpha d_1 q(qn+1) - q^2 \\ &= q[an + (\alpha d_1 - q)] = pq.\end{aligned}$$

The Legendre symbol

$$(76)\quad (-ad \mid p) = (-1)^{(p-1)/2}(\alpha \mid p) = (-1)^{[(p-1)/2]\,[(\alpha+1)/2]}(pq \mid \alpha)(q \mid \alpha)$$
$$= (-1)^{[(p-1)/2]\,[(\alpha+1)/2]+1+(\alpha-1)/2} = (-1)^{[(p+1)/2]\,[(\alpha+1)/2]} = 1,$$

since $\alpha + 1 \equiv 0 \pmod 8$. Hence there exists an integer η such that

$$(77)\qquad -ad \equiv \eta^2 \pmod p.$$

It is evident that if we choose r to satisfy $ar \equiv \eta \pmod p$ (77) becomes

$$(78)\qquad d + ar^2 \equiv 0 \pmod p.$$

Also $d + ar^2 \equiv 0 \pmod q$ since $d \equiv a \equiv 0 \pmod q$. But p is prime to q. Hence $d + ar^2 \equiv 0 \pmod{ab - t^2}$. We have proved

LEMMA 11. *Let $a = \alpha d$, $\alpha \equiv 7 \pmod 8$, $(q \mid \alpha) = -1$ for a prime factor q of d. Then a is a coefficient of $\Sigma(d)$.*

We have now shown that if $a = \sigma$ is any integer with no square factor which is not of the form $\sigma = \alpha d$, $\alpha \equiv 7 \pmod 8$, $(q \mid \alpha) = 1$ for every prime factor q of d, a is a coefficient of $\Sigma(d)$. We may now easily show conversely that no integer of the above form is a coefficient of $\Sigma(d)$. For let *

$$(79)\qquad \phi = fx^2 + by^2 + \cdots$$

be a form of determinant d and let

$$(80)\qquad \Phi = AX^2 + BY^2 + CZ^2 + \cdots$$

be the *reciprocal form*. It is well known † that f and C may be so chosen that the three integers f, C, $2d$ are relatively prime in pairs. Also †

$$(81)\qquad fC\Phi = \lambda^2 + C\mu^2 + fd\nu^2.$$

Suppose first that $(C \mid p) = -(-1 \mid p)$ for some factor p of d and let a be an integer of the above form. Then if a is represented by (79) we have

$$(82)\qquad fCa = \lambda_0^2 + C\mu_0^2 + fd\nu_0^2,$$

---

* It is the following part of the proof of Theorem 9 that is due to the referee and replaces my original proof which was a direct application of Theorem 1, and so did not use any of the theory of reciprocal forms, simultaneous representation, or generic invariants.

† Cf. L. E. Dickson's *Studies in the Theory of Numbers* for the above elementary properties of reciprocal forms. For the above properties of the invariant Ψ, see H. J. S. Smith's *Collected Papers*, pp. 455-507, in particular p. 464 and p. 473. In our work $d = D\Delta^2$ is an odd integer with no square factor so that $\Delta = 1$, D is odd and the integers α and β of Smith are $+1$.

for integers λ_0, μ_0, ν_0. But $a \equiv 0 \pmod{d}$ and $d \equiv 0 \pmod{p}$. Hence $\lambda_0^2 + C\mu_0^2 \equiv 0 \pmod{p}$. If $\mu_0 \not\equiv 0 \pmod{p}$ then $C \equiv -\epsilon^2 \pmod{p}$, $(C \mid p) = (-1 \mid p)$, a contradiction. Hence $\mu_0 \equiv 0 \pmod{p}$ so that $\lambda_0 \equiv 0 \pmod{p}$ and $f(Ca - d\nu_0^2) \equiv 0 \pmod{p^2}$. But f is prime to $2d$ so that $(C\alpha - \nu_0^2)d \equiv 0 \pmod{p^2}$. Thus $C\alpha \equiv \nu_0^2 \pmod{p}$ since $\alpha d = a$, $d = pd_1$ is a product of distinct primes. Then

$$(83) \qquad (C\alpha \mid p) = 1, \quad (C \mid p) = (\alpha \mid p) = -(-1 \mid p),$$

so that, since $(p \mid \alpha) = 1$ by hypothesis,

$$\begin{aligned}(84) \quad (\alpha \mid p) &= (-1)^{[(p-1)/2]\,[(\alpha-1)/2]}(p \mid \alpha) \\ &= (-1)^{[(p-1)/2]\,[(\alpha-1)/2]} = -(-1 \mid p) = (-1)^{[(p-1)/2]+1}.\end{aligned}$$

Hence $(-1)^{[(p-1)/2]\,[(\alpha+1)/2]} = -1$ which is false since $\alpha + 1 \equiv 0 \pmod{8}$.

Hence if a is represented by a form (79) of determinant d it must be so that $(C \mid p) = (-1 \mid p)$ for every prime factor p of d, whence $(C \mid d) = (-1 \mid d)$. The integer

$$(85) \qquad \Psi = (-1)^{\frac{1}{2}(df+1)\frac{1}{2}(F+1)}$$

is an invariant * of ϕ, that is has the same value for every integer f represented by ϕ and F represented simultaneously by Φ. By this we mean that if we pass to a form equivalent to ϕ and simultaneously to the corresponding new reciprocal form then Ψ has the same value when we substitute for f the new f and for C the new C. But it is also known * that

$$(86) \qquad \Psi(C \mid d) = (-1)^{\frac{1}{2}(d+1)}$$

for our case where d has no square factor and is odd.* Since $(C \mid d) = (-1 \mid d)$ then $\Psi = (-1)^{(d+1)/2+(d-1)/2} = (-1)^d = -1$. But a is a coefficient of $\Sigma(d)$, appears as leading coefficient of a form equivalent to ϕ and hence $\Psi = (-1)^{(da+1)/2\,.\,(F+1)/2} = 1$ since

$$da + 1 = \alpha d^2 + 1 \equiv \alpha + 1 \equiv 0 \pmod{8},$$

a contradiction.

We have proved

THEOREM 9. *Let a and d have no square factors, d be odd. Then a is a coefficient of $\Sigma(d)$ if and only if a is not an integer of the form*

$$(87) \qquad a = \alpha d, \quad \alpha \equiv 7 \pmod{8}, \quad (p \mid \alpha) = 1$$

for every prime factor p of d.

---

* *Loc. cit.*

6. *Integers represented by sets* $\Sigma(d)$. Let $a = \rho^2\sigma$, $d = \gamma^2\delta$ where σ and δ have no square factor. If a is represented by $\Sigma(d)$ then, by Theorem 6, σ is a coefficient of $\Sigma(d)$. But then we shall show that σ has not the form

$$(88) \qquad \sigma = \alpha\delta, \quad \alpha \equiv 7 \pmod 8, \quad \alpha \text{ prime to } d, \quad (p \mid \alpha) = 1$$

for every prime factor p of d.

For otherwise let σ have the above form. By Theorem 8 there exists a factor Δ of d such that $\alpha\Delta = \sigma_1$ has no square factor and is a coefficient of $\Sigma(\Delta)$. But this is impossible if Δ is odd by Theorem 9 with $d = \Delta$. Hence $\Delta = 2\Delta_0$, $\sigma_1 = 2\sigma_0$. By Theorem 2, σ_0 is a coefficient of $\Sigma(\Delta_0)$ which is again false by Theorem 9.

Conversely let σ have not the above form. If δ is odd and σ is either not divisible by α or $\sigma = \alpha\delta$ with $\alpha \not\equiv 7 \pmod 8$ or $\alpha \equiv 7 \pmod 8$ but $(p \mid \alpha) = -1$ for some prime factor p of δ then, by Theorem 9, σ is a coefficient of $\Sigma(\delta)$ and, by Theorem 3, of $\Sigma(d)$. Hence let $\sigma = \alpha\delta$, $\alpha \equiv 7 \pmod 8$, $(p \mid \alpha) = 1$ for every prime factor p of δ while we still are considering the case δ odd.

Suppose α is not prime to d. Since α is prime to δ then $\alpha = \alpha_1 p$ while p is a prime divisor of d and hence of γ but not of δ. Also α is odd so that p is odd. Write $\delta_0 = p\delta$, which is an integer with no square factor, $\sigma_0 = \alpha_1\delta$. The integer σ_0 is not divisible by δ_0 so that, by Theorem 9, σ_0 is a coefficient of $\Sigma(\delta_0)$. By Theorem 7, $\sigma = \sigma_0 p$ is a coefficient of $\Sigma(p\delta_0)$. But $p\delta_0 = p^2\delta$, $\gamma = \gamma_0 p$, $d = \gamma_0^2 p^2\delta = \gamma_0^2(p\delta_0)$ and, by Theorem 3, σ is a coefficient of $\Sigma(d)$.

Finally let α be prime to d but let there exist a prime factor p of d such that $(p \mid \alpha) = -1$. By our above hypothesis p must divide γ not δ. Hence $\gamma = \gamma_0 p$ and if $\delta_0 = \delta p$ then $\sigma p = \alpha\delta_0$ is a coefficient of $\Sigma(\delta_0)$ by Theorem 9. By Theorem 7, σp^2 is a coefficient of $\Sigma(p\delta_0)$, $p\delta_0 = p^2\delta$. By Theorem 4, σ is a coefficient of $\Sigma(p^2\delta)$ so that, by Theorem 3, σ is a coefficient of $\Sigma(\gamma_0^2p^2\delta) = \Sigma(d)$.

We have now proved that if δ is odd and σ is not of the form (88) then σ is a coefficient of $\Sigma(d)$. Let then δ be even, $\delta = 2\delta_1$, $d = 2d_1$, $d_1 = \gamma^2\delta_1$. If σ is odd and not of the form (88) then 2σ has not the form (88) with d replaced by d_1, δ by δ_1. But then, as we have proved, 2σ is a coefficient of $\Sigma(d_1)$. By Theorem 2, 4σ is a coefficient of $\Sigma(d)$ and, by Theorem 3, of $\Sigma(4d)$. Then by Theorem 4, σ is a coefficient of $\Sigma(d)$. Hence let $\sigma = 2\sigma_1$. Then σ_1 has not the form (88) for d replaced by d_1 so that σ_1 is a coefficient of $\Sigma(d_1)$. By Theorem 2, σ is a coefficient of $\Sigma(d)$.

We have therefore proved

THEOREM 10. *Write any two positive integers in their unique forms*

(89) $$d = \gamma^2\delta, \quad a = \rho^2\sigma,$$

where δ and σ have no square factors. An integer a is represented by a set $\Sigma(d)$ if and only if σ is not an integer of the form

(90) $$\sigma = \alpha\delta, \quad \alpha \equiv 7 \pmod 8,$$

such that α is prime to d and

(91) $$(p \mid \alpha) = 1$$

for every prime factor p of d.

7. *The regularity of sets $\Sigma(d)$.* L. E. Dickson has called a ternary quadratic form regular if it represents exclusively all the integers not in a certain set, given by a finite number of formulae,* of arithmetic progressions. We shall prove that every $\Sigma(d)$ is regular in the above sense and in fact

THEOREM 11. *Let $p_1, \cdots, p_r$ be the distinct odd prime factors of d, $P = p_1 \cdot p_2 \cdots p_r$, $d = \gamma^2\delta$ where δ has no square factor. Let η_i range over all the finite number of least residues of p_i which satisfy the condition*

(92) $$(\eta_i \mid p_i) = (-1 \mid p_i),$$

and let A_ϵ range over the corresponding finite number of least solutions of the finite system (as the η_i vary) of sets of congruences

C_ϵ: $$x \equiv 7 \pmod 8, \quad x \equiv \eta_i \pmod{p_i}.$$

Then the set $\Sigma(d)$ represents exclusively all positive integers not of the form

(93) $$4^k p_1^{2k_1} \cdots p_r^{2k_r} (8nP + A_\epsilon)\delta,$$

and is regular.

For we need only show that the condition

(94) $$a = \rho^2\alpha\delta, \quad \alpha \text{ prime to } d, \quad \alpha \equiv 7 \pmod 8, \quad (p \mid \alpha) = 1$$

for every prime factor p of d is equivalent to a of the form (93).

If a has the form (93) then we may write $\rho = 2^k \prod\limits_i p_i^{k_i}\nu$ where ν is odd and prime to d and need only show that an integer $A = \nu^2\alpha$, α with no square factor, has the property $\alpha \equiv 7 \pmod 8$, $(p/\alpha) = 1$, α prime to d if and only if $A = 8nP + A_\epsilon$.

If $A = \nu^2\alpha$ then $A = \nu^2\alpha \equiv \alpha \equiv 7 \pmod 8$. Also A is prime to d and

* That is, apart from square factors as in $4^k(8n + 7)$, actually a finite number of arithmetic progressions $t^2(an + b)$ where t is quite arbitrary.

$(A \mid p) = (\alpha \mid p) = (-1)^{[(p-1)/2]\,[(\alpha-1)/2]}(p \mid \alpha) = (-1 \mid p)$ since $(p \mid \alpha) = 1$, $\alpha \equiv 7 \pmod 8$. Hence A is a solution of a set C_ϵ and, by the Chinese remainder Theorem, $A \equiv A_\epsilon \pmod{8P}$, $A = 8Pn + A_\epsilon$. Conversely let $A = 8Pn + A_\epsilon$. Then $A \equiv A_\epsilon \equiv 7 \pmod 8$, $A \equiv \eta \pmod p$ for every prime factor p of d, where $(\eta \mid p) = (-1 \mid p)$. But then A is prime to d, A is odd. Write $A = \nu^2\alpha$ where α has no square factor. Then ν is odd $\nu^2 \equiv 1 \pmod 8$, $\alpha \equiv A \equiv 7 \pmod 8$ and $(p \mid \alpha) = (-1)^{[(p-1)/2]\,[(\alpha-1)/2]}(\alpha \mid p) = (-1 \mid p)^2 = 1$ as desired and Theorem 11 is proved.

We may take $A_\epsilon = -1$ as a particular instance since $-1 \equiv 7 \pmod 8$ and obviously satisfies the requirement $A_\epsilon = \eta_i \pmod{p_i}$. Moreover every integer $8nd - 1$ is an integer of the form $8nP - 1$ since P is a factor of d, $d = mp$, $8nd - 1 = 8(nm)P - 1$. We have therefore proved

THEOREM 12. *Every set $\Sigma(d)$, $d = \gamma^2\delta$, where δ has no square factor represents no integer of the form $\delta(8nd - 1)$ and hence no $\Sigma(d)$ represents all positive integers.*

This theorem, by the way, provides a new proof of a well known result

THEOREM 13. *No positive ternary quadratic form represents all positive integers.*

8. *Sets of positive n-aries.* Consider an n-ary quadratic form

$$\phi(x_1, \cdots, x_n) = \sum_{i,j}^{1,\ldots,n} a_{ij}x_ix_j, \tag{95}$$

where $n \geq 4$, the integers a_{ij} are such that

$$A = (a_{ij}) \tag{96}$$

is a symmetric matrix. The form (95) is called positive if $\phi(x_1, \cdots, x_n) \geq 0$ for all integers $x_1, \cdots, x_n$ and $\phi(x_1, \cdots, x_n) = 0$ only when $x_1 = \cdots = x_n = 0$. The determinant d of the matrix A is called the determinant of the form (95) and, when ϕ is positive d is a positive integer. We shall consider in particular forms

$$\phi(x_1, \cdots, x_n) = f(x_1, x_2, x_3) + \sum_{i=4}^{n} x_i^2, \tag{97}$$

where $f(x_1, x_2, x_3)$ is a positive ternary of determinant d so that (97) has determinant d and is positive.

Consider the set $\Sigma(n, d)$ of all positive n-aries of determinant d. Evidently $\Sigma(n, d)$ contains forms (97). Let $a = \gamma^2\delta$ where δ has no square factor. Similarly let $a = \rho^2\sigma$. If σ is not divisible by δ then there exists a

positive ternary $f(x_1, x_2, x_3)$ of determinant d and integers α, β, γ such that $a = f(\alpha, \beta, \gamma)$ by Theorem 10. But then the n-ary (97) in $\Sigma(d)$ represents a, that is $\phi(\alpha, \beta, \gamma, 0, \cdots, 0) = a$. Next let $\sigma = \alpha\delta$ be divisible by δ. Then $\rho^2(\sigma - 1) = \rho^2\rho_1{}^2\sigma_0 = \rho_0{}^2\sigma_0$ where σ_0 is not divisible by δ since evidently σ_0 divides $\alpha\delta - 1$. By Theorem 10 there exists a positive ternary $f(x_1, x_2, x_3)$ of determinant d representing $\rho^2(\sigma - 1)$. Then there exist integers α, β, γ for which $f(\alpha, \beta, \gamma) = \rho^2\sigma - \rho^2 = a - \rho^2$. Then if ϕ is the form (97) we have $\phi(\alpha, \beta, \gamma, \rho, 0, \cdots, 0) = a - \rho^2 + \rho^2 = a$. We have therefore proved, *using only the property that a set* $\Sigma(3, d)$ *represents all integers* $a = \rho^2\sigma$, σ *not divisible by* δ.

THEOREM 14. *Every set* $\Sigma(n, d)$ *of all positive n-ary quadratic forms of determinant* d, $n \geqq 4$, *represents all positive integers.*

THE UNIVERSITY OF CHICAGO,
CHICAGO, ILLINOIS.

ON REPRESENTATION OF INTEGERS BY INDEFINITE TERNARY QUADRATIC FORMS OF QUADRATFREI DETERMINANT.*

By Arnold E. Ross.†

1. The object of this paper is a study of representation of integers by indefinite ternary quadratic forms whose determinant is free from square factors.

A. Meyer ‡ gave conditions under which an indefinite form f with relatively prime invariants Ω, Δ (Δ odd, $\Omega \not\equiv 0 \pmod 4$)) represents an odd integer m prime to Ω and not divisible by certain prime factors of Δ. Dickson § extended these conditions to the case when only m or both m and Δ are double an odd integer, retaining all other restrictions of Meyer.

Employing a method of Dirichlet ¶ and Markoff's ‖ table of indefinite ternary forms Dickson ** studied representation of integers by forms of negative determinant $-D$, where $D \leqq 83$ and was either a prime, double a prime, the product of two distinct primes, or double such a product.

In this paper we generalize the above mentioned method of Dirichlet and employ this generalization and a theorem of Meyer †† to solve the problem of representation of integers by indefinite ternary quadratic forms whose determinant is odd and is free from square factors.

We shall prove the following

Theorem 1.‡‡ *Let f be an indefinite ternary quadratic form whose*

* Read before the American Mathematical Society, November, 1929.

† National Research Fellow.

‡ A. Meyer, *Vierteljahrsschrift Naturf. Gesellschaft Zürich*, Vol. 29 (1884), pp. 209-222.

§ L. E. Dickson, *Studies in the Theory of Numbers*, Ch. V, Chicago, 1930.

¶ G. L. Dirichlet, *Journal für Mathematik*, Vol. 40 (1850), pp. 228-232.

‖ A. W. Markoff, *Mém. Imp. Acad. Sc. St. Petersbourg*, series 8, Vol. 23 (1909), No. 7, 22 pp. For an extension of this table see L. E. Dickson, *Ibid.*, pp. 150-151.

** For the details of Professor Dickson's method see S. Silberfarb, "Representation by indefinite ternary quadratic forms," *Dissertation, University of Chicago*, 1929. Also R. H. Marquis, "The representation of integers," *Dissertation, University of Chicago*, 1929.

†† A. Meyer, *Journal für Mathematik*, Vol. 108 (1891), p. 139. See also L. E. Dickson, *Ibid.*, p. 54, Theorem 47.

‡‡ The last mentioned results of Dickson, Silberfarb, and Marquis are all instances of this theorem and its companion theorem for even determinants. For this latter see A. E. Ross, *Proceedings of the National Academy of Sciences*, Vol. 18 (1932), pp. 600-608, § 1.

determinant is odd and free from square factors. Let Ω and Δ be the invariants of f. Then $\Omega = \pm 1$. Write $\rho_1, \cdots, \rho_\alpha$ *for those odd prime divisors of Δ for which*

$$(1.1) \qquad (F \mid \rho_i) = -(-\Omega \mid \rho_i) \qquad (i = 1, \cdots, \alpha),$$

and let $\pi_1, \cdots, \pi_v$ be those odd prime divisors of Δ for which

$$(1.2) \qquad (F \mid \pi_j) = (-\Omega \mid \pi_j) \qquad (j = 1, \cdots, v).$$

Then if α is even f represents every integer a of none of the types

$$(1.3) \qquad \rho_i^{2k+1}[n\rho_i + \mu(-\Delta_i, \rho_i)]$$

and no integers of any of these types. Here $n = 0, \pm 1, \pm 2, \cdots k = 0, 1, 2, \cdots, \Delta = \rho_i \Delta_i$ *and*

$$(1.4) \qquad \begin{array}{l} \text{for a given integer } x \text{ prime to } p \\ \mu(x, p) \text{ runs over all the least residues} \\ \text{of } p \text{ satisfying } (\mu(x,p) \mid p) = (x \mid p). \end{array}$$

If α is odd, f represents every integer a of none of the types (1.3) *and*

$$(1.5) \qquad 4^k(8n - \Delta)$$

where $k = 0, 1, 2, \cdots$, *and no integers of any of these types.*

The above theorem shows that in the case of an indefinite ternary quadratic form f whose determinant is odd and free from square factors, all integers not represented by f form several families of arithmetical progressions depending in a simple way on the prime factors of the determinant and the generic characters of f. The same has been found to be true for indefinite ternary quadratic forms in a much more general * case than the one considered here.

It is of interest to note that in what follows our lemmas, including the extension of the Dirichlet method, apply to positive as well as to indefinite forms.

2. *Integers not represented by f.* In this section we study the relation between the generic characters of f and integers not represented by f.

LEMMA 1. *Let f be a primitive ternary quadratic (definite or indefinite) form with invariants Ω and Δ. Let δ be a prime divisor of Δ for which*

$$(2.11) \qquad (F \mid \delta) = -(-\Omega \mid \delta).$$

---

* These results will appear in a forthcoming paper. For a brief report see A. E. Ross, *ibid.*, §§ 4, 5, and 7.

Further write $\Delta = \delta\Delta_1$, *and let* δ *be prime to* $\Omega\Delta_1$. *Then* f *does not represent integers of the form*

$$\delta^{2k+1}[n\delta + \mu(-\Delta_1, \delta)], \tag{2.12}$$

where $\mu(-\Delta_1, \delta)$ *runs over all those least residues of* δ *whose quadratic character is the same as that of* $-\Delta_1$, *i. e.*

$$(\mu(-\Delta_1, \delta) \mid \delta) = (-\Delta_1 \mid \delta). \tag{2.13}$$

Write

$$f = ax^2 + by^2 + cz^2 + 2ryz + 2sxz + 2txy, \tag{2.14}$$

and let

$$F = Ax^2 + By^2 + Cz^2 + 2Ryz + 2Sxz + 2Txy \tag{2.15}$$

be the reciprocal of f. Then *

$$aCf = CX^2 + \Omega Y^2 + \Omega\Delta a z^2 \tag{2.16}$$

where

$$X = ax + ty + sz, \qquad Y = Cy - Rz. \tag{2.17}$$

We may assume † that a and C are relatively prime and have no odd prime factor in common with $\Omega\Delta$.

If m is represented by f, then (2.16) holds with f replaced by m. Let $m = m_1\delta$. Then

$$aCm_1\delta = CX^2 + \Omega Y^2 + \Omega\Delta a z^2, \tag{2.18}$$

whence

$$CX^2 \equiv -\Omega Y^2 \pmod{\delta}$$

and in view of (2.11) and (2.15), $X \equiv Y \equiv 0 \pmod{\delta}$. Write $X = \delta X_1$, $Y = \delta Y_1$. Substituting into (2.18) and dividing through by the common factor δ we obtain $aCm_1 \equiv \Omega\Delta_1 a z^2 \pmod{\delta}$, whence, since a is prime to δ,

$$Cm_1 \equiv \Omega\Delta_1 z^2 \pmod{\delta}. \tag{2.19}$$

If m_1 is prime to δ, then $(m_1 \mid \delta) = (C\Omega\Delta_1 \mid \delta)$. Since by (2.11), $(C\Omega \mid \delta) = -(-1 \mid \delta)$, we have $(m_1 \mid \delta) = -(-\Delta_1 \mid \delta)$. Therefore if $(m_1 \mid \delta) = (-\Delta_1 \mid \delta)$, m is not represented by f. This proves our lemma for the case of $k = 1$.

To complete the proof we note that if f does not represent l, it does not

---

* See A. E. Ross, *ibid.*, § 5.

† H. J. S. Smith, *Collected Mathematical Papers*, Vol. 1 (1894), §§ 5 and 9; L. E. Dickson, *Studies in the Theory of Numbers*, pp. 15-17, Chicago, 1930; P. Bachman, *Die Arithmetik der Quadratischen Formen v.* 1, p. 64.

represent $\delta^2 l$. For if we let $m = \delta^2 l$, then $m_1 \equiv 0 \pmod{\delta}$ and by (2.19), δ divides z. Thus δ divides all of X, Y, z and hence by (2.17) also all of x, y, z, and therefore f represents l, a contradiction.

LEMMA 2. *Let f be a properly primitive ternary quadratic (definite or indefinite) form whose determinant is odd and is free from square factors. If (i) f is definite and*

$$(F \mid \Delta) = (-\Omega \mid \Delta), \tag{2.21}$$

or if (ii) f is indefinite and

$$(F \mid \Delta) = -(-\Omega \mid \Delta), \tag{2.22}$$

then f does not represent integers of the type

$$4^k(8n - \Delta). \tag{2.23}$$

Since f and F are both properly primitive we may without loss of generality assume further that a and C are both odd.* We shall first prove our lemma in case $\Omega > 0$, whence $\Omega = 1$. Then Δ is positive or negative according as f is positive or indefinite. Smith's character condition † in this case becomes

$$\Psi = (F \mid \Delta)(-1)^{\frac{1}{2}(\Delta+1)}$$

where

$$\Psi = (-1)^{\frac{1}{2}(\Omega C+1)\frac{1}{2}(\Delta a+1)} \tag{2.24}$$

Write $\Delta = e \mid \Delta \mid$. Then $\Psi = (F \mid \Delta)(-e)(-1 \mid \Delta)$. Therefore if $\Omega = 1$, in both cases (i) and (ii) of our lemma

$$\Psi = -1. \tag{2.25}$$

Relations (2.24) and (2.25) imply

$$\Omega C \equiv 1, \quad \Delta a \equiv 1 \pmod{4}. \tag{2.26}$$

Hence, by (2.16),

$$aCf \equiv C(X^2 + Y^2 + z^2) \pmod{4}. \tag{2.27}$$

If $f \equiv -\Delta_1$, (mod 8), then $aCf \equiv -C \pmod{4}$, whence $X^2 + Y^2 + z^2 \equiv 3 \pmod{4}$, and therefore $X \equiv Y \equiv z \equiv 1 \pmod{2}$. Then $-aC\Delta \equiv C + \Omega + \Omega\Delta a \pmod{8}$. Hence

---

* In fact to be able to do that we choose $C \equiv \Omega \pmod{4}$ and hence the first one of congruences (2.26) will hold in any case. See L. E. Dickson, *Studies*, p. 15.

† H. J. S. Smith, *Collected Mathematical Papers*, Vol. 1 (1894), p. 470.

$$0 \equiv (a\Delta + 1)C + (a\Delta + 1)\Omega = (C + \Omega)(a\Delta + 1) \pmod 8.$$

But $a\Delta \equiv 1 \pmod 4$, and therefore $C \equiv -\Omega \pmod 4$ contrary to (2.261).

If $f = 4m \equiv 0 \pmod 4$, then $X^2 + Y^2 + z^2 \equiv 0 \pmod 4$ by (2.27). Therefore $X \equiv Y \equiv z \equiv 0 \pmod 2$, whence $x \equiv y \equiv z \equiv 0 \pmod 2$ by (2.17), and f represents m. This proves Lemma 2 for $\Omega = 1$.

To complete the proof we need only note that every ternary quadratic form with $\Omega = -1$ is a negative * of one with $\Omega = 1$ and that if $f_1 = -f$, then $\Omega_1 F_1 = \Omega F$, $\Omega_1 C_1 = \Omega C$, and $\Delta_1 a_1 = \Delta a$.

3. *Dirichlet construction.* We consider a ternary quadratic (definite or indefinite) form f whose determinant is odd and free from square factors. Let a be an integer of none of the types not represented by f in view of Lemmas 1 and 2. We shall in this section generalize a method of Dirichlet † to prove that every such integer a is represented by some form ϕ in the same genus as f.

Adopting the already mentioned conventions for the sign of Ω, we shall assume in what follows that $\Omega > 0$ which in our case implies that $\Omega = 1$. As we have pointed out in the closing paragraph of the preceding section this assumption does not essentially restrict the generality. Then as in the proof of Lemma 2, Δ is positive or negative according as f is positive or indefinite.

If an integer $\gamma^2 a$ is not of the type (2.12) or (2.23) (or, what is the same, (1.3) or (1.5)), then a is not of that type, and therefore we need to prove the desired result only for integers a without square factors.

LEMMA 3. *Let f be a ternary quadratic (positive or indefinite) form whose determinant is odd and free from square factors. Let $\Omega = 1$ and Δ be the invariants of f. Let $\rho_1, \cdots, \rho_a$ be all of those positive odd prime divisors of Δ for which* (1.1) *holds and write $\pi_1, \cdots, \pi_v$ for those for which* (1.2) *holds.*

(*i*) *Let f be positive. Then if α is odd and a is positive, quadratfrei, and is of none of the types* (1.3) *or if α is even and a is positive quadratfrei, and is of none of the types* (1.3) *and* (1.5), *there exists a positive form*

$$\phi = ax^2 + by^2 + cz^2 + 2ryz + 2sxz \tag{3.11}$$

of the same genus as f, having a as its leading coefficient and hence representing a properly.

---

* For adopted conventions for the sign of Ω, see H. J. S. Smith, *Collected Mathematical Papers*, Vol. 1 (1894), p. 456, or L. E. Dickson, *Studies*, p. 10.

† G. L. Dirichlet, *Journal für Mathematik*, Vol. 40 (1850), pp. 228-232; E. Laudau, *Vorlesungen über Zahlentheorie*, B. 1 (1927), pp. 123-125.

(*ii*) *Let f be indefinite. Then if α is even and a is quadratfrei and is of none of the types* (1.3) *or if α is odd and a is quadratfrei and is of none of the types* (1.3) *and* (1.5), *there exists an indefinite form ϕ as in* (3.11), *of the same genus as f and which has a as the leading coefficient and hence represents a properly.*

We wish to show thus that for every integer a subject to the restrictions of our lemma we can choose integers b, c, r, s so that the form (3.11) has for its invariants $\Omega = 1$ and Δ, i.e.,

$$aA - bs^2 = \Delta, \quad \text{here} \quad A = bc - r^2, \tag{3.12}$$

and so that its generic characters coincide with those of f.

We shall proceed with the construction.

Write $\rho_1 \cdots \rho_a = R$, $\pi_1, \cdots, \pi_v = E$. Let $T = (a, R)$, $S = (a, E)$ and write

$$R = TP, \quad E = SQ, \quad a = TSa_1, \tag{3.13}$$

$$\Delta = eD \quad \text{where} \quad D = |\Delta| > 0. \tag{3.14}$$

Then

$$D = RE = PQST. \tag{3.15}$$

Since by the assumption a is not of the form (1.3),

$$(a/\tau \mid \tau) = -(-\Delta/\tau \mid \tau) = -(-eD/\tau \mid \tau)$$

for every odd prime divisor τ of T. Hence by (3.13_3) and (3.15)

$$(a_1 \mid \tau) = -(-ePQ \mid \tau). \tag{3.16}$$

Write p, q for odd prime divisors of P and Q respectively, and let π be a power of 2. Take

$$s = T, \quad A = \pi Sd, \quad d = 8PQTu + v, \tag{3.17}$$

$$v \equiv S \pmod 8, \quad \pi Sa_1 v \equiv ePQ \pmod T, \tag{3.18}$$

$$(v \mid p) = -(-\pi S \mid p), \quad (v \mid q) = (-S\pi \mid q). \tag{3.19}$$

Write

$$\pi Sa_1 v - ePQ = T2^\lambda w, \tag{3.21}$$

where w is odd. In view of (3.12) let

$$T^2 b = aA - eD. \tag{3.22}$$

Then by (3.13) and (3.17)

$$T^2b = T^2S[8\pi a_1PQSu + 2^\lambda w] = T^2S2^\lambda b_1,$$

where

(3.23) $$2^\lambda b_1 = 8M + 2^\lambda w, \quad M = \pi a_1PQS.$$

By (3.21) every common divisor of πSa_1 and $2^\lambda w$ divides also PQ, similarly every common divisor of PQ and w must divide πSa_1v. Since πSa_1v is prime to PQ each one of πSa_1 and PQ and hence their product M is prime to $2^\lambda w$.

Next, by (3.21), (3.13$_3$), and (3.18$_1$),

(3.24) $$2^\lambda w \equiv \pi aS - ePQT \pmod 8.$$

1°. If

(3.25) $$\pi a_1 \equiv 0 \pmod 2,$$

then $\lambda = 0$ and b_1 in (3.23) is odd.

2°. Assume next that

(3.26) $$\pi a_1 \equiv 1 \pmod 2, \quad \text{whence} \quad \pi = 1.$$

Then (3.24) becomes $2^\lambda w \equiv aS - ePQT \pmod 8$, and hence $2^\lambda w \equiv 0 \pmod 4$ if and only if $aS \equiv ePQT \pmod 4$, that is if and only if $a \equiv ePQST \pmod 4$. Therefore if

(3.27) $$a \equiv -ePQST \equiv -\Delta \pmod 4,$$

then $\lambda = 1$ and b_1 in (3.23) is odd.

By (3.23) we may write

(3.28) $$b_1 = \sigma Mu + w,$$

where $\sigma = 8$ or 4 according as $\lambda = 0$ or 1. Then for every a_1 and π which satisfy (3.25) or (3.26) and (3.27), b_1 in (3.28) is an odd integer for every integral value of u. Moreover, the coefficients σM and w of the arithmetical progression $\sigma Mu + w$ are relatively prime by the above. Hence by the Dirichlet * theorem on primes in an arithmetical progression, there are infinitely many primes of the form (3.28) and hence we may assume that b_1 is a positive odd prime not dividing Δ. If also

(3.29) $$(-A \mid b_1) = 1,$$

then there exists an integer r such that $-A \equiv r^2 \pmod{b_1}$. Since b_1 is odd and prime to Δ we may choose r odd and $\equiv 0 \pmod S$.

* G. L. Dirichlet, *Abhandlungen der Königlichen Akademie der Wissenschaften*, pp. 108-110, Berlin, 1837. E. Laudau, *Vorlesungen über Zahlentheorie*, B. 1 (1927), pp. 79-96.

Write $-A = r^2 - b_1c_1$. Then by (3.17$_2$) and the choice of r, $c_1 \equiv 0 \pmod{S}$, and if $\pi = 1$, c_1 is even. We may write therefore $c_1 = 2^\lambda Sc$. Then

$$A = b_1 2^\lambda Sc - r^2 = bc - r^2.$$

The form ϕ thus determined has determinant Δ. The adjoint Φ of ϕ is properly primitive since the g. c. d. of $A = bc - r^2$ and $B = ac - s^2$ is prime to Δ by (3.17) and (3.18$_2$). Hence $|\Omega| = 1$. Since b is positive, ϕ is positive if a and Δ are positive, for then the two upper left hand corner principal minors a and ab of the determinant of ϕ and also the determinant itself are positive. If Δ is negative, then ϕ is indefinite since it represents a positive integer b and has a negative determinant.* Hence, in accord with our conventions, $\Omega = 1$. The generic characters of ϕ with respect to the odd prime factors σ, p, q of S, P, Q respectively, are

$$(\Phi \mid \sigma) = (B \mid \sigma) = (ac - s^2 \mid \sigma) = (-T^2 \mid \sigma) = (-1 \mid \sigma),$$
$$(\Phi \mid p) = (A \mid p) = (\pi Sv \mid p) = -(-1 \mid p),$$
$$(\Phi \mid q) = (A \mid q) = (\pi Sv \mid q) = (-1 \mid q),$$

by (3.13$_3$), (3.17), and (3.19). Since by (3.16) and (3.18$_2$)

$$(3.291) \qquad (v \mid \tau) = (\pi S a_1 v \mid \tau)(\pi S \mid \tau)(a_1 \mid \tau) = -(-\pi S \mid \tau),$$

the characters of ϕ with respect to every odd prime factor τ of T are

$$(\Phi \mid \tau) = (A \mid \tau) = (\pi Sv \mid \tau) = -(-1 \mid \tau).$$

Hence if (3.29) holds, the constructed form ϕ is the one desired in Lemma 3.

To complete the proof of our lemma it remains thus to show that for every integer a satisfying conditions of our lemma we can choose π so that either (3.25) or (3.26) and (3.27) hold and so that (3.29) is true.

For every a_1 and π satisfying (3.25) or (3.26) and (3.27), and for every integral value of u we have, in view of (3.17), (3.18), and (3.22),

$$(-A \mid b_1) = (-\pi \mid b_1)(Sd \mid b_1)$$
$$(Sd \mid b_1) = (b_1 \mid Sd) = (T^2 2^\lambda b_1 \mid Sd) = (-ePQT \mid Sd)$$
$$= (PQT \mid Sd) = (Sd \mid PQT) = (Sv \mid PQT).$$

By (3.19) and (3.291)

$$(v \mid PQT) = (-1)^a(-\pi S \mid PQT).$$

Therefore

$$(3.31) \quad (-A \mid b_1) = (-1)^a(-1 \mid b_1)(-1 \mid PQT)(\pi \mid b_1)(\pi \mid PQT).$$

* See L. E. Dickson, *Studies*, p. 10, § 7. One can verify these assertions directly by multiplying (2.16) by Ω and replacing f by ϕ and ΩC by ab.

Let, (i_1), f be positive and α be odd or, (ii_1), f be indefinite and α be even. Then for every * a not of the types (1.3) we take $\pi = 4$. Then

$$b_1 \equiv w \equiv -ePQT \pmod 4$$

by (3.28) and (3.21). Since $e = +1$ or -1 according as f is positive or indefinite, (3.29) holds in both of the above cases by (3.31).

Next, let (i_2), f be positive and α be even or, (ii_2), f be indefinite and α be odd. Let a be an integer of none of the types (1.3) and (1.5).

If

$$a \not\equiv -\Delta \equiv -ePQST \pmod 4 \tag{3.32}$$

we take $\pi = 2$. Then

$$b_1 \equiv w \pmod 8 \tag{3.33}$$

by (3.28). If a is even, whence a is double an odd,

$$w \equiv -ePQT + 4 \pmod 8 \tag{3.34}$$

by (3.21). Remembering again that $e = +1$ or -1 according as f is positive or indefinite, we see that (3.29) holds by (3.31). Next let a be odd. Then $a \equiv ePQST \pmod 4$ by (3.32) and

$$w \equiv -ePQT + 2ePQT \equiv ePQT \pmod 8.$$

Then $b_1 \equiv ePQT$ and (3.31) implies (3.29).

If $a \equiv -\Delta \pmod 4$, then

$$a \equiv -ePQST + 4 \pmod 8$$

since a is not of the form (1.5). In this case we take $\pi = 1$. Then $b_1 \equiv w \pmod 4$. Also, by (3.21) and (3.35), $2w \equiv -2ePQT + 4 \pmod 8$, whence $w \equiv -ePQT + 2 \pmod 4$. Then

$$b_1 \equiv -ePQT + 2 \pmod 4$$

and again (3.31) implies (3.29).

4. *Proof of Theorem* 1. Lemmas 1 and 2 show that no integer of the form (1.3) or (1.3) and (1.5) according as α is even or odd, are represented by an indefinite form f satisfying the conditions of Theorem 1. If $\Omega = 1$, Lemma 3 shows that every integer not excluded by these lemmas is represented by some form ϕ in the genus of f. But by Meyer's † theorem

* If f is positive we need only to consider positive integers a.

† A. Meyer, *Journal für Mathematik*, Vol. 108 (1891), p. 139; L. E. Dickson, *Studies*, p. 54.

there is but one class in every genus of indefinite ternary quadratic forms of determinant wihch is free from square factors. Hence ϕ is equivalent to f, and a is represented by f as well as by ϕ. If $\Omega = -1$ we augment Lemma 3 by an argument similar to that at the end of Section 2.

In case f is a positive ternary form of odd quadratfrei determinant, and if the genus of f contains but one class,* Lemmas 1, 2, and 3 permit us to write down at once the families of the arithmetical progressions giving the totality of integers not represented by f.

For example consider forms

$$f_1 = x^2 + 2y^2 + 3z^2 - 2yz$$

and

$$f_2 = 2x^2 + 2y^2 + 3z^2 + 2yz + 2xz + 2xy.$$

Determinants of f_1 and f_2 are equal to 5 and 7 respectively. Their generic characters are

$$(F_1 \mid 5) = (2 \mid 5) = -1 = -(-\Omega \mid 5) = (F_1 \mid \Delta),$$
$$(F_2 \mid 7) = (3 \mid 7) = -1 = (-\Omega \mid 7) = (F_1 \mid \Delta).$$

Their genera contain but one class each. (See L. E. Dickson, *Studies in the Theory of Numbers*, p. 181, Chicago, 1930.) Applying Lemmas 1 and 3 to f_1 we see that f_1 represents all integers save those of the form $5^{2k+1}(5n+1)$ or $5^{2k+1}(5n+4)$. Similarly Lemmas 2 and 3 show that f_2 represents all integers not of the form $4^k(8n+1)$.

California Institute of Technology.

---

* Representation of integers by a genus of positive ternary forms has been studied by Jones (see B. W. Jones, *Transactions of the American Mathematical Society*, Vol. 33 (1931), pp. 92-110, 111-124). However his final results involve the auxiliary parameters α, β, γ of a lemma of Smith (H. J. S. Smith, *Collected Mathematical Papers*, Vol. 1 (1894), p. 460).

ON THE STRÖMGREN-WINTNER NATURAL TERMINATION PRINCIPLE.

By G. Baley Price.

Introduction. The subject of analytic continuation of periodic orbits, developed by Poincaré,† has been considered in two papers recently by Wintner.‡ In the first of these he proved the Strömgren-Wintner Natural Termination Principle for groups of periodic orbits. The only example of the principle which has been given so far is the restricted problem of three bodies, but it is so complicated that the groups of periodic orbits can be studied only by numerical integration of the equations of motion. The present paper furnishes a simple example which can be treated mathematically. Also it gives illustrations of Poincaré's theorems on the disappearance of periodic orbits by pairs † and on the change of stability of periodic orbits.§

Wintner shows that Poincaré's theorem on the disappearance of periodic orbits by pairs is without significance in the study of groups of periodic orbits. By means of the present example it is pointed out that there are other points of view, or other problems, in dynamics in which the disappearance of periodic orbits in Poincaré's sense is a real and significant phenomenon.

Finally, this paper adds further information about a class of dynamical systems previously investigated.¶ In order to make the example as simple and specific as possible, a special case of the general class is studied here, but the results hold for any of the systems for which $r_x^2 + u_x^2 \neq 0$ [see (8) below], and only obvious modifications in the results are necessary in case this condition does not hold.

1. *The equations of motion.* We shall consider the motion of a heavy particle on a surface of revolution S of genus one. Choose the positive ζ-axis directed downward [R 1, p. 753 (i. e., reference 1 at the end of this paper)].

† Poincaré, *Méthodes Nouvelles de la Mécanique Céleste*, Vol. 1 (1892), chapter 3.

‡ Wintner, "Beweis des E. Strömgrenschen dynamischen Abschlussprinzips der periodischen Bahngruppen im restringierten Dreikörperproblem," *Mathematische Zeitschrift*, Vol. 34 (1931), pp. 321-349; "Sortengenealogie, Hekubakomplex und Gruppenfortsetzung," *Mathematische Zeitschrift*, Vol. 34 (1931), pp. 350-402.

§ Poincaré, *Méthodes Nouvelles de la Mécanique Céleste*, Vol. 3 (1899), pp. 343-351.

¶ Price, "A class of dynamical systems on surfaces of revolution," *American Journal of Mathematics*, Vol. 54 (1932), pp. 753-768.

Then $u(x) = g\zeta(x)$, where g is the acceleration of gravity, and from the first paper [R 1, equations (9), (12), (13)] we have

$$r^2y' = c, \qquad \text{[the integral of areas]} \tag{1}$$

$$x'' = (c^2r_x + gr^3\zeta_x)/r^3, \tag{2}$$

$$x'^2 = [2r^2(g\zeta + h) - c^2]/r^2. \tag{3}$$

Here h is the energy constant. The functions v and w are [R 1, (14), (15)]

$$v = 2r^2(g\zeta + h), \tag{4}$$

$$w = -gr^3\zeta_x/r_x, \qquad r_x \neq 0. \tag{5}$$

Finally, using (4) and (5), we can write (2) and (3) in the form

$$x'' = r_x(c^2 - w)/r^3, \qquad r_x \neq 0 \tag{6}$$

$$x'^2 = (v - c^2)/r^2. \tag{7}$$

The following relation is important also [R 1, (5)]:

$$r_x^2 + \zeta_x^2 = 1. \tag{8}$$

2. *Properties of v and w.* A direct computation gives

$$v_x = 2[2rr_x(g\zeta + h) + gr^2\zeta_x], \tag{9}$$

which can be written in the form

$$v_x = 2r_x(v - w)/r, \qquad r_x \neq 0. \tag{10}$$

Now since (8) holds, we see from (9) that v_x never vanishes when $r_x = 0$. Then from (10) we obtain the proof of the following lemma.

LEMMA 1. *The derivative v_x vanishes when and only when $(v - w)$ vanishes.*

The following lemma states another important fact.

LEMMA 2. *A necessary and sufficient condition that the parallel $x = x^*$ be a trajectory is that $v - c^2 = 0$, $v_x = 0$ on this parallel.*

A necessary and sufficient condition that $x = x^*$ be a trajectory is that $x' = 0$, $x'' = 0$ on $x = x^*$. Then since $v_x = 0$ when and only when $(v - w) = 0$, the lemma follows from (6) and (7).

Let a parallel $x = x^*$ on which $v_x = 0$ be designated by P^*. On a parallel P^*, $r_x \neq 0$ and $v = w$, and we find that

$$v_{xx} = -2r_xw_x/r. \tag{11}$$

Hence, on a parallel P^*, v has a maximum if $r_x w_x > 0$ and a minimum if $r_x w_x < 0$. Using this result, we can prove without difficulty the following lemma. The details are left to the reader.

LEMMA 3. *A necessary and sufficient condition that v have a maximum (minimum) on a parallel $P^*: x = x^*$ is that $x = x^*$ be an interior point of some interval in which $r_x w_x \geqq 0$ $(r_x w_x \leqq 0)$.*

One further lemma is necessary.

LEMMA 4. *A necessary and sufficient condition that v have a point of inflection with a horizontal tangent on a parallel $P^*: x = x^*$ is that $r_x w_x$ have opposite signs in sufficiently small intervals on opposite sides of $x = x^*$.*

The condition is necessary, for if $r_x w_x$ has the same sign on the two sides of x^*, then v has a maximum or minimum by lemma 3. Also, the condition is sufficient, for by lemma 3 v can have neither a maximum nor a minimum; hence, it has a point of inflection with a horizontal tangent.

Now plot $v = v(x)$ and $w = w(x)$ on the same field of rectangular coördinates. Since v and w are periodic with period ω [R 1, (2)], we may restrict attention to the interval $0 \leqq x < \omega$. At each zero of r_x, w has a vertical asymptote. Now r has at least one maximum and one minimum, at which r_x vanishes and changes sign. Because of (8) then, w has at least two vertical asymptotes at which w is asymptotic to one end of the asymptote on one side and to the other end on the other side.

Now w is fixed by the choice of the surface S and does not vary for a given system. On the other hand, v varies with the energy constant h. But since v is finite for all values of x and h, the curves v and w have a certain number of intersections. By lemma 1, $v_x = 0$ at each point of intersection. The nature of v at the point of intersection is further determined by lemmas 3 and 4.

If v and w intersect at a point where $w_x \neq 0$, or at a point where w has a point of inflection with a horizontal tangent, then v has a maximum or minimum by lemma 3. At such a point v and w cross. If v and w intersect at a point where w has a maximum or minimum, then v has a point of inflection with a horizontal tangent. At such a point v and w do not cross.

3. *Groups of periodic orbits.* It is possible to choose h uniquely so that v intersects w at an arbitrary point of w. Furthermore, by lemma 2 any intersection of v and w at which v and w are positive corresponds to a parallel P^* which is a closed periodic orbit on S. As h varies, the intersections of v and w vary, and in general analytically. Thus the closed orbits P^* can be

continued analytically with h to form what Wintner [R 2] has called a *group of periodic orbits.*

The points of the curve w for which $w > 0$ are in one-to-one continuous correspondence with closed periodic orbits P^* in (x, y, h) space. Then each connected piece of w lying in the region $w > 0$ is in one-to-one continuous correspondence with the orbits of a group. We may therefore refer to a connected piece of w in the region $w > 0$ as the *graph of a group.*

Let us investigate the manner in which these groups terminate. In the present case, a group terminates in one of two ways. In the first place, the graph may have an end point on the line $w = 0$. As we approach such an end point along the graph, the period of the corresponding orbit P^* becomes infinite [see (1) and lemma 2] with h and the dimensions of the orbit remaining finite. In the second place, the graph may have a vertical asymptote. As we approach such an asymptote along the graph of the group, the energy constant h becomes positively infinite for the corresponding orbit P^*. The period of the orbit approaches zero, and its dimensions remain finite.

These results are in accord with the Natural Termination Principle. The example does not show a group which closes into itself, nor a group which terminates because the dimensions of the orbit become infinite.

Let us view this example in the light of Poincaré's conclusions [R 3, Vol. 1, p. 83]. Consider any intersection of v and w. As h varies, this intersection varies and generates what we may call a *branch of a group of periodic orbits.* Start at any point on the graph of a group and continue along the graph in each direction as far as possible without passing a maximum or minimum of $w = w(x)$; the periodic orbits which correspond to any such piece of the graph form what we call a branch of a group. There is at most one periodic orbit in each branch of a group for a given value of h. Poincaré's conclusion was that a branch of a group can be continued with increasing and decreasing h unless it combines with a second such branch and disappears. Our example shows exactly how this may happen.

Consider a maximum of $w = w(x)$. For a suitable value of h, v will intersect w twice in the neighborhood of this maximum; at one intersection v has a maximum and at the other a minimum (see lemma 3). As h increases, the maximum and minimum vary until v is finally tangent to w at the maximum of w. The maximum and minimum of v have combined to form a point of inflection with a horizontal tangent. At the same time the two corresponding periodic orbits P^* have combined to form a multiple orbit. For still larger values of h, there is no periodic orbit P^* in the neighborhood. The orbits have disappeared as described by Poincaré.

Furthermore, it should be observed that of the two orbits which combine and disappear, one is stable and the other is unstable [see (11) and R 1, (20)]. This result is in agreement with one of Poincaré's well known theorems [R 3, Vol. 3, pp. 343-351]. We see [(11) above and R 1, (20)] that the characteristic exponents which Poincaré calls α and $-\alpha$ are zero when and only when $w_x = 0$. As we follow along the graph of a group, there is no change in stability when we pass a point of inflection with a horizontal tangent. On the other hand, there is a change in stability whenever we pass a maximum or minimum on the graph of the group. From Poincaré's point of view, a stable and an unstable orbit combine and disappear with proper variation of h at these points.

Whenever we consider the totality of orbits for a given value of the energy constant h [it has been customary to study dynamical systems from this point of view in the past; see R 4, p. 270], Poincaré's theorem on the disappearance of periodic orbits by pairs will be meaningful and significant. In particular, suppose the present problem is being studied by means of a surface of section [R 1, § 4]. In setting up a surface of section we must consider the totality of orbits for a given value of h. The members of a group of periodic orbits which exist for the given value of h give rise to fixed points in the surface transformation on the surface of section. As h varies, these fixed points vary and sometimes appear and disappear by pairs. Whenever two branches of a group combine and disappear with variation of h, two fixed points combine and disappear. The fixed points of a surface transformation are highly significant features of the transformation. Thus, although Poincaré's theorem on the disappearance of periodic orbits by pairs is without significance in the study of groups, it is both meaningful and significant in certain other problems.

The group as considered by Wintner is obtained as follows. Take any branch of a group; it may be that at one end, or both, this branch joins to other branches. Join on these branches, and then join any branches that have an end in common with these branches, and so on. In the present problem we are stopped only when we reach a branch which terminates on the line $w = 0$, or which has a vertical asymptote. The group appears as the totality of branches that can be reached by continuation from a single branch.

The development of the problem of analytic continuation of periodic orbits can be sketched briefly as follows: Poincaré considered the branches of a group rather than the group itself, and showed that a branch might terminate by combining with a second branch [R 3, Vol. 1, p. 83]. He considered no other possibilities, however. Later he recognized that a branch might terminate because the period becomes infinite [R 5, p. 258]. Birkhoff

states [R 5, p. 258]: "To make possible an extension to a preassigned interval $\mu_0 \leqq \mu \leqq \mu_1$ it is necessary to prove that the period of the varying periodic orbit does not become infinite." The question of the sufficiency of the condition is not considered explicitly. Finally, Strömgren and Wintner have emphasized the point of view of the group, and Wintner has proved that a group does not terminate so long as the period, energy, and dimensions of the orbit remain finite.

Finally, we may emphasize that the Natural Termination Principle does not prove that there is at least one member of a given group for every value of the parameter h. In the above example it is true that there exists a periodic orbit P^* in each of at least two groups for all values of h for which motion over the entire surface is possible, but this is not true for every group. We shall show a group to illustrate this fact.

Choose the surface S so that

$$\begin{aligned} &r_x > 0, && x_1 \leqq x \leqq x_2 \\ &\zeta_x(x_1) = \zeta_x(x_2) = 0, && \\ &\zeta_x < 0, && x_1 < x < x_2. \end{aligned}$$

Then $w(x_1) = w(x_2) = 0$, and w is positive and finite for $x_1 < x < x_2$. We thus have a group which terminates at both ends because the period becomes infinite. For suitable small values of h, there are certain periodic orbits of this group in the system for the corresponding value of h. As h increases, these orbits combine and disappear by pairs. For h sufficiently large, there is no periodic orbit of the group in the system.

UNION COLLEGE,
SCHENECTADY, NEW YORK.

REFERENCES

[1] Price, "A class of dynamical systems on surfaces of revolution," *American Journal of Mathematics*, Vol. 54 (1932), pp. 753-768.

[2] Wintner, "Beweis des E. Strömgrenschen dynamischen Abschlussprinzips der periodischen Bahngruppen im restringierten Dreikörperproblem," *Mathematische Zeitschrift*, Vol. 34 (1931), pp. 321-349.

[3] Poincaré, *Méthodes Nouvelles de la Mécanique Céleste*, Vol. 1 (1892), Vol. 3 (1899).

[4] Birkhoff, "The restricted problem of three bodies," *Rendiconti del Circolo Matematico di Palermo*, Vol. 39 (1915), pp. 265-334.

[5] Birkhoff, "Dynamical systems with two degrees of freedom," *Transactions of the American Mathematical Society*, Vol. 18 (1917), pp. 199-300.

UPON A STATISTICAL METHOD IN THE THEORY OF DIOPHANTINE APPROXIMATIONS.

By Aurel Wintner.

Introduction.

Let

$$f(s) = \sum_{n=1}^{\infty} a_n \exp(\lambda_n s); \qquad a_n \neq 0$$

denote a Dirichlet series possessing linearly independent real exponents λ_n and a domain (i. e. half-plane or strip) in which $f(s)$ is absolutely convergent. Let α be a real number in the interior of this domain and set

$$z = z(t) = x(t) + iy(t) = f(\alpha + it)$$

where $-\infty < t < +\infty$. The values taken by $z(t)$ are, according to Jessen,* distributed asymptotically in such a way that there exists, in the (x, y)-plane, a continuous function $D = D(x, y)$ determining the density of this distribution, i. e. the density of probability (relative frequency as $t \to \infty$) of the values taken by $z(t) = x(t) + iy(t)$. The method of Jessen is built, on the one hand, upon an integration theory in a space of infinitely many dimensions and, on the other hand, upon the Kronecker-Weyl approximation theorem.

In the present paper the treatment of the distribution problem belonging to the almost-periodic function $z(t)$ will be based upon the general statistical or momentum method, as developed, for the one-dimensional case, by the author,† and recently extended to higher spaces by Haviland.‡ It will be proven that the continuous density function D, the existence of which (i. e. Jessen's result) need not be presupposed, is related to the distribution function § ρ belonging to the real part $x(t)$ of $z(t)$ by an integral equation of the Abel type. Since ρ is explicitly known ¶ we thus obtain an analytical method

* B. Jessen, *Bidrag til Integraltheorien for Funktioner of uendelig mange Variable*, Copenhagen, 1930.

† A. Wintner, "Diophantische Approximationen und Hermitesche Matrizen. I.," *Mathematische Zeitschrift*, Vol. 30 (1929), pp. 290-319 (more particularly pp. 310-311). This paper will be referred to as *I*.

‡ E. K. Haviland, "On statistical methods in the theory of almost-periodic functions," *Proceedings of the National Academy of Sciences*, Vol. 19 (1933), May issue.

§ First introduced *loc. cit. I*.

¶ A. Wintner, "On an application of diophantine approximation to the repartition problems of dynamics," *Journal of the London Mathematical Society*, Vol. 7 (1932),

for an effective control of D. With the use of Bessel functions, the application of this explicit method yields the result that $D(x, y)$ not only is everywhere continuous but also possesses derivatives of arbitrarily high order save at most at the origin $x = y = 0$, without being analytic im grossen. The question as to whether D is analytic im kleinen remains open. On the other hand, the method works just as well in the "non-analytic" case,* where the series $f(s)$ is absolutely convergent not in a domain (i. e. half-plane or strip) but only on the isolated line $s = \alpha + it$. Hence we start directly with an arbitrary almost-periodic function

$$(1) \qquad z(t) = x(t) + iy(t) = \sum_{j=1}^{\infty} r_j \exp i\lambda_j(t - t_j); \qquad r_j > 0$$

$(-\infty < t < +\infty)$ where the frequencies λ_j are supposed to be linearly independent, in which case, according to a theorem of Bohr,† of necessity

$$(2) \qquad R < +\infty \quad \text{where} \quad R = \sum_{j=1}^{\infty} r_j.$$

It may be mentioned that the ultimate reason for the occurance of the Abel integral equation reducing D to ρ lies in the fact that on account of the Laplace-Fourier transforms of D and ρ this reduction is a transformation of "planes waves" into "spherical waves."

Applications to the μ-function of Lindelöf will be given in a subsequent paper.

The Distribution of the Real Component.

The distribution function $\rho = \rho(\xi)$ of an arbitrary ‡ real-valued almost-periodic function $x(t)$ is defined for $-\infty < \xi < +\infty$ as

$$(3) \qquad \lim_{T=+\infty} \operatorname{meas} \{x(t) \leqq \xi;\ T\}/2T$$

where $\{x(t) \leqq \xi;\ T\}$ denotes the set of all those points t for which both inequalities $x(t) \leqq \xi$, $|t| < T$ are satisfied, and $\operatorname{meas}\{x(t) \leqq \xi;\ T\}$ is the Lebesgue measure § of this set. The limit (3) exists ¶ save for a denumerable

pp. 242-246. This paper will be referred to as *II*. Cf. also "Ueber die statistische Unabhängigkeit," *Mathematische Zeitschrift*, Vol. 36 (1933), pp. 618-629. This paper will be referred to as *III*.

* In reality the question regarding the analytic continuation of such a function $f(\alpha + it)$ does not seem to have been treated yet in the literature.

† H. Bohr, "Zur Theorie der fast-periodischen Funktionen. I.," *Acta Mathematica*, Vol. 45 (1925), p. 103.

‡ The linear independence of the frequencies is not yet supposed.

§ This is at present a Jordan content inasmuch as $x(t)$ is almost-periodic and therefore continuous.

¶ *Loc. cit. I.*

set of exceptional values $\xi = \xi_m$ which, if they exist, are always discontinuity points * of the monotone function $\rho(\xi)$. The latter is defined as the limit (3) if $\xi \neq \xi_m$ and as the arithmetical mean of $\rho(\xi + 0)$ and $\rho(\xi - 0)$ if $\xi = \xi_m$. An exceptional point ξ_m may actually exist.† On the other hand, it is possible that ξ_m is a discontinuity point of $\rho(\xi)$ without being ‡ an exceptional point ξ_m.

Now let $x(t)$ be the real part of (1), i. e. suppose that the frequencies of the almost-periodic function $x(t)$ are linearly independent. Then $\rho(\xi)$ is everywhere continuous §; hence (3) exists for every ξ. We shall see later on that all derivatives of $\rho(\xi)$ exist. Let $\rho_k(\xi)$ denote the distribution function belonging to the partial sum

(4) $$x_k(t) = \sum_{j=1}^{k} r_j \cos \lambda_j (t - t_j)$$

of

(5) $$x(t) = \sum_{j=1}^{\infty} r_j \cos \lambda_j (t - t_j); \qquad \sum_{j=1}^{\infty} r_j = R < + \infty.$$

Then ¶

(6) $$\rho_{k+1}(\xi) = \int_{-\infty}^{+\infty} \rho_k(\xi - \eta)\, d\sigma_{k+1}(\eta); \qquad \rho_1(\xi) = \sigma_1(\xi),$$

where $\sigma_j(\xi)$ denotes the distribution function belonging to the periodic function

(7) $$a_j(t) = r_j \cos \lambda_j (t - t_j);$$

i. e.

---

* *Loc. cit. I.*

† H. Bohr, "Kleinere Beiträge zur Theorie der fastperiodischen Funktionen. II.," *Det Kgl. Danske Videnskabernes Selskab. Meddelelser*, Vol. 10, No. 10 (1930).

‡ For let the continuous function $x(t)$ be periodic with the period 1 and let it be of bounded variation in the fundamental region $0 \leqq t \leqq 1$. Suppose further that $x(t)$ is zero when $|t - n| \leqq 1/4$, $n = 0, \pm 1, \pm 2, \cdots$ but that $x(t) \neq 0$ for all other values of t. Since the Fourier partial sum $x_k(t)$ is a periodic trigonometric polynomial, its distribution function $\rho_k(\xi)$ is everywhere continuous. Furthermore, $x_k(t)$ approaches the limit $x(t)$ uniformly when $k \to \infty$. Finally, the limit (3) exists for every ξ inasmuch as $x(t)$ is periodic. The limit $\rho(\xi)$ of $\rho_k(\xi)$ possesses, however, a discontinuity at $\xi = 0$.

§ A. Wintner, "Ueber die Stetigkeit der asymptotischen Verteilungsfunktion bei inkommensurablen Partialschwingungen," *Mathematische Zeitschrift*, Vol. 37 (1933), not yet appeared.

¶ *Loc. cit. II.* The recursion formula (6) yields a k-fold iterated Stieltjes integral for $\rho_{k+1}(\xi)$, viz.

$$\rho_{k+1}(\xi) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \sigma_1(\xi - \eta_1)\, d\sigma_2(\eta_1 - \eta_2) \cdots d\sigma_k(\eta_{k-1} - \eta_k)\, d\sigma_{k+1}(\eta_k).$$

This detailed representation takes the place of the shortened expression (18) in the paper *II*, a formula whose meaning is obvious from (19), *loc. cit. II.*

$$\text{(8)} \qquad \begin{cases} \sigma_j(\xi) = 0 \quad \text{for} \quad -\infty < \xi < -r_j \\ \sigma_j(\xi) = 1 - [\arccos(\xi/r_j)]/\pi \quad \text{for} \quad -r_j \leqq \xi \leqq r_j \\ \sigma_j(\xi) = 1 \quad \text{for} \quad r_j < \xi < +\infty \end{cases}$$

where $0 \leqq \arccos \leqq \pi$. Furthermore,

$$\text{(9)} \qquad \rho(\xi) = \lim_{k=\infty} \rho_k(\xi)$$

holds for those * values of ξ which are continuity points of $\rho(\xi)$; hence (9) holds for all values of ξ. Finally,†

$$\text{(10)} \qquad L(s;\, \rho_k) = \prod_{j=1}^{k} L(s;\, \sigma_j)$$

where $L(s;\, \nu)$ denotes the Laplace-Fourier transform

$$\text{(11)} \qquad L(s;\, \nu) = \int_{-\infty}^{+\infty} \exp(is\xi)\, d\nu(\xi)$$

of the typical distribution function $\nu(\xi)$ and s is an arbitrary real or complex parameter. Since $|x_k(t)|$ and $|x(t)|$ are, according to (4) and (2), not larger than R, it follows from the definition (3) of a distribution function that

$$\rho_k(\xi) = 0 \quad \text{for} \quad -\infty < \xi < -R, \qquad \rho_k(\xi) = 1 \quad \text{for} \quad R < \xi < +\infty$$

and

$$\text{(12)} \quad \rho(\xi) = 0 \quad \text{for} \quad -\infty < \xi < -R, \qquad \rho(\xi) = 1 \quad \text{for} \quad R < \xi < +\infty.$$

Accordingly, all Stieltjes integrations $\int_{-\infty}^{+\infty}$ may be replaced by $\int_{-R}^{R}$. Hence, from (9) and (11),

$$\text{(13)} \qquad \lim_{k=\infty} L(s;\, \rho_k) = L(s;\, \rho)$$

by virtue of the Helly theorem on term-by-term integration.‡ On comparing (10) with (13) there results the multiplicative relation §

---

* *Loc. cit. II.*

† *Loc. cit. II.* Cf. G. Doetsch, "Die Integrodifferentialgleichungen vom Faltungstypus," *Mathematische Annalen*, Vol. 89 (1923), pp. 192-207.

‡ E. Helly, "Ueber lineare Funktionaloperationen," *Sitzungsberichte der mathematisch-naturwissenschaftlichen Klasse der Kaiserl. Akademie der Wissenschaften zu Wien*, Vol. 121 (1912), pp. 265-297.

§ The existence of the infinite product (14) is for all values of s assured by (13). Since $J_0(0) = 1$, there follows from (15) and (2) by Schwarz's Lemma a finer result, viz. the uniform convergence of the series

$$\sum_{j=1}^{\infty} |L(s;\, \sigma_j) - 1|$$

in every fixed s-circle. Similar remarks hold regarding the infinite products occurring later on.

$$L(s;\ \rho) = \prod_{j=1}^{\infty} L(s;\ \sigma_j) \tag{14}$$

expressing the statistical independence * of the distributions σ_j belonging to the partial vibrations (7) of (5).

From (11) and (8) we have

$$L(s;\ \sigma_j) = (1/\pi) \int_{-r_j}^{r_j} (r_j^2 - \xi^2)^{-1/2} \exp(is\xi)\, d\xi,$$

i. e.

$$L(s;\ \sigma_j) = (2/\pi) \int_0^1 (1 - \xi^2)^{-1/2} \cos(sr_j\xi)\, d\xi,$$

or, on placing $\xi = \cos\theta$,

$$L(s;\ \sigma_j) = (2/\pi) \int_0^{\pi/2} \cos(sr_j \cos\theta)\, d\theta.$$

Hence †

$$L(s;\ \sigma_j) = J_0(r_j s). \tag{15}$$

From (11), (14) and (15) there results

$$L(s;\ \rho) = \int_{-\infty}^{+\infty} \exp(is\xi)\, d\rho(\xi) = \prod_{j=1}^{\infty} J_0(r_j s). \tag{16}$$

We notice here that the distribution of $x(t)$ is symmetric with respect to the origin, i. e.

$$\rho(\xi) + \rho(-\xi) = 1. \tag{17}$$

On account of (9) it is sufficient to prove that

$$\rho_k(\xi) + \rho_k(-\xi) = 1 \tag{18}$$

holds for every k. Now from (8)

$$\sigma_j(\xi) + \sigma_j(-\xi) = 1. \tag{19}$$

Hence (18) holds for $k = 1$ inasmuch as $\rho_1 = \sigma_1$. Suppose that (18) holds for a fixed value of k. Since from (6)

$$\rho_{k+1}(-\xi) = \int_{-\infty}^{+\infty} \rho_k(-\xi - \eta)\, d\sigma_{k+1}(\eta) = -\int_{-\infty}^{+\infty} \rho_k(-\xi + \zeta)\, d\sigma_{k+1}(-\zeta),$$

where $\eta = -\zeta$, there results from (18) and (19) the equality

* Cf. F. Hausdorff, "Beitraege zur Wahrscheinlichkeitsrechnung," *Berichte über die Verhandlungen der Königl. Saechsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-physikalische Klasse*, Vol. 53 (1901), pp. 152-178. This paper discusses also the general methods in Calculus of Probability, which have a connection with the present problem.

† R. Courant und D. Hilbert, *Methoden der mathematischen Physik, I.*, 1924, p. 393.

$$\rho_{k+1}(-\xi) = -\int_{-\infty}^{+\infty}[1-\rho_k(\xi-\zeta)]\,d[1-\sigma_{k+1}(\zeta)],$$

i. e.

$$\rho_{k+1}(-\xi) = \int_{-\infty}^{+\infty} d\sigma_{k+1}(\zeta) - \int_{-\infty}^{+\infty}\rho_k(\xi-\zeta)\,d\sigma_{k+1}(\zeta),$$

or, by virtue of (6),

$$\rho_{k+1}(\xi) + \rho_{k+1}(-\xi) = \int_{-\infty}^{+\infty} d\sigma_{k+1}(\zeta) = 1,$$

inasmuch as the last integral represents the total variation of a monotone function (8). Hence (18) holds for every k. From (16) and (17) there results

$$(20) \qquad L(s;\,\rho) = 2\int_0^{+\infty}\cos(s\eta)\,d\rho(\eta) = \prod_{j=1}^{\infty} J_0(r_j s);$$

hence, by virtue of (12),

$$(21) \qquad 2\int_0^{+\infty}\sin(s\eta)/\eta\,d\rho(\eta) = \int_0^{s}\prod_{j=1}^{\infty} J_0(r_j l)\,dl.$$

For positive values of the independent variable we need the appraisals

$$(22) \qquad \Big|\prod_{j=1}^{\infty} J_0(r_j\eta)\Big| < \Gamma_m/\eta^m; \qquad (m = 0, 1, 2, \cdots),$$

where Γ_m is a constant depending upon m but independent of $\eta > 0$. First, the well-known asymptotic formula *

$$J_0(\eta) \sim \eta^{-\frac{1}{2}}(2/\pi)^{\frac{1}{2}}\cos(\eta - \pi/4); \qquad \eta \to +\infty$$

assures the existence of a constant C for which

$$|J_0(\eta)| < C/\eta^{\frac{1}{2}}.$$

Accordingly,

$$\Big|\prod_{j=1}^{2m} J_0(r_j\eta)\Big| < \Gamma_m/\eta^m,$$

where $\Gamma_m = C^{2m}/(r_1 r_2 \cdots r_{2m-1} r_{2m})^{\frac{1}{2}}$. Hence (22) is obvious inasmuch as

$$J_0(X) = \int_0^{2\pi}\cos(X\cos\theta)\,d\theta/2\pi$$

has, for real values of X, a modulus $\leqq 1$.

We now restrict s in (21) to real and non-negative values and write ξ instead of s. Thus

$$2\int_0^{+\infty}\sin(\xi\eta)/\eta\,d\rho(\eta) = \int_0^{\xi}\prod_{j=1}^{\infty} J_0(r_j\eta)\,d\eta; \qquad \xi \geqq 0.$$

* Courant-Hilbert, *op. cit.*, p. 435.

This integral equation for the monotone continuous function ρ may be solved, by virtue of (12), by means of the Gauss-Fourier inversion formula which yields *

$$(23)\qquad \rho(\xi)-\rho(0)=(1/\pi)\int_0^{+\infty}\{[\sin(\xi\eta)/\eta]\prod_{j=1}^{\infty}J_0(r_j\eta)\}\,d\eta$$

for $\xi \geqq 0$. It is clear from (17) that (23) holds for $\xi < 0$ also. The expression

$$(1/\pi)\int_0^{+\infty}(d^k\{\cdot\ \cdot\ \cdot\}/d\xi^k)\,d\eta$$

resulting from (23) by k-fold formal differentiation is, by virtue of (22), absolutely and uniformly convergent for $-\infty < \xi < +\infty$. In order to see this, it is sufficient to choose $m \geqq k+2$. Since m, and therefore k, may be chosen arbitrarily large, it follows † that *the distribution function* $\rho(\xi)$ *possesses for* $-\infty < \xi < +\infty$ *derivatives of arbitrarily high order.*

Hence from (17)

$$(24)\qquad \rho(0)=\tfrac{1}{2},\qquad \rho^{(k)}(0)=0;\qquad\qquad (k=2,4,6,\cdot\cdot\cdot).$$

Similarly from (12)

$$(25)\qquad \rho^{(k)}(R)=0,\qquad \rho^{(k)}(-R)=0;\qquad\qquad (k=1,2,3,\cdot\cdot\cdot),$$

although $\rho(\xi)$ is known ‡ to be nowhere constant in the range $-R \leqq \xi \leqq R$. Thus the behavior of $\rho(\xi)$ at $\xi = \pm R$ is the same as that of Cauchy's example

$$\exp(-1/\xi^2)$$

at $\xi = 0$.

Let us notice that the distribution function $\rho_k(\xi)$ belonging to the finite sum (4) cannot possess derivatives of arbitrarily high order if k has a fixed value. Correspondingly, infinitely many appraisals (22) break down if the infinite product is replaced by a finite one.

First, $\rho_1 = \sigma_1$ is everywhere continuous, its derivative is, however, infinite at $\xi = \pm r_1$. The function ρ_2 has been considered by Bessel in his celebrated

* The validity of the Gauss-Fourier inversion formula (cf. F. Hausdorff, *loc. cit.*), which is at present (23), is assured under conditions which are essentially more general than (12). Cf., for instance, T. C. Burkill, "The expression in Stieltjes integrals of the inversion formulae of Fourier and Hankel," *Proceedings of the London Mathematical Society*, Series 2, Vol. 25 (1926), pp. 513-524.

† Cf., for instance, E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, Second Edition, Vol. II, p. 359, Cambridge University Press, 1926.

‡ *Loc. cit. I.* It follows that the function $x(t)$ takes on every value between $-R$ and R. The latter fact is contained in the Kronecker approximation theorem also.

paper on the Gaussian frequency curve.* The first derivative of ρ_2 is a complete elliptic integral of the first kind and is infinite at the four points $\xi = \pm (r_1 + r_2)$, $\xi = \pm (r_1 - r_2)$, two of which coincide when $r_1 = r_2$. The function ρ_3 possesses everywhere a continuous first derivative but the second derivative is infinite at some points, and so on, so that ρ_k is the smoother, the farther we go in Bessel's statistical † iteration process (6).

It is clear from (2) that the limit function ρ cannot be related to the Gaussian frequency curve.‡

The Markoff condition for the validity of the Gauss law § takes in our case the form

$$\lim_{k=\infty} S_{2n}(k) : S_2(k) = 0, \quad (n = 2, 3, \cdots), \text{ where } S_n(k) = \left(\sum_{j=1}^{k} r_j{}^n\right)^{1/n}.$$

This condition is, however, not a necessary one (Liapounoff).

An Integral Equation for the Central Waves.

For later purposes (cf. p. 327) we consider in the present chapter a function $\delta(r)$ implicitly defined for $0 \leqq r \leqq R$ as a continuous solution of the functional equation

$$(26) \qquad \rho'(r) = 2\int_r^R (q^2 - r^2)^{-\frac{1}{2}} q\delta(q)\,dq; \qquad 0 \leqq r \leqq R.$$

There exists exactly one such function and it possesses, save at the origin $r = 0$, derivatives of arbitrarily high order. Furthermore,

$$(27) \qquad \delta^{(k)}(R) = 0, \qquad (k = 0, 1, 2, \cdots).$$

Finally,

$$(28) \qquad \int_0^{2\pi}\int_0^R r\delta(r)\exp\{i(u\cos\vartheta + v\sin\vartheta)r\}dr\,d\vartheta = L(\{u^2 + v^2\}^{\frac{1}{2}};\ \rho)$$

where u and v are arbitrary real or complex parameters.

In order to prove these statements we first reduce (26) to Abel's integral equation

* F. W. Bessel, *Abhandlungen*, Vol. 2 (1876), pp. 378-380.

† Cf. also H. Bohr and B. Jessen, "Om Sandsynlighedsfordelinger ved Addition af konvekse Kurven," *Det Kgl. Danske Videnskabernes Selskabs Skrifter*, Series 8, Vol. 12 (1929), No. 3.

‡ Cf. in this connection F. Hausdorff, *loc. cit.*

§ Cf. R. Deltheil, *Erreurs et moindres carrés*, Paris, 1930, pp. 71-74; M. Fréchet and J. Shohat, "A proof of the generalized central limit theorem in the theory of probability," *Transactions of the Mathematical Society*, Vol. 33 (1931), pp. 533-543.

(29) $$\chi(X) = \int_0^X (X-Y)^{-\frac{1}{2}} \tau(Y)\,dY; \qquad 0 \leqq X \leqq R^2$$
by placing
(30) $$X = R^2 - r^2, \qquad Y = R^2 - q^2$$
and
(31) $$\chi(X) = \rho'(\sqrt{R^2 - X}), \qquad \tau(X) = \delta(\sqrt{R^2 - X})$$
(hence χ is given and τ is the unknown function). Since $\rho(\xi)$ has for every ξ derivatives of any order, the function $\chi(X)$ possesses, according to (31), derivatives of arbitrarily high order in the half-open range $0 \leqq X < R^2$; furthermore, by virtue of (24), (25) and (31),

(32) $$\chi^{(k)}(0) = 0, \qquad (k = 0, 1, 2, \cdots),$$

and the first derivative $\chi'(X)$ exists and is continuous in the closed range $0 \leqq X \leqq R^2$. Hence * (29) has exactly one continuous solution τ in this closed range, viz. the one represented by Abel's inversion formula

(33) $$\tau(X) = \int_0^X (X-Y)^{-\frac{1}{2}} \chi'(Y)\,dY/\pi; \qquad 0 \leqq X \leqq R^2.$$

On combining (30) and (31) with (33) we see that (26) possesses the unique continuous solution

(34) $$\delta(r) = -\int_r^R (q^2 - r^2)^{-\frac{1}{2}} \rho''(q)\,dq/\pi; \qquad 0 \leqq r \leqq R.$$

We have now to prove that in the half-open range $r \neq 0$ all derivatives of $\delta(r)$ exist and satisfy the relations (27). In other words [cf. (30), (31)], we have to prove that in the half-open range $0 \leqq X < R^2$ the function $\tau(X)$ possesses derivatives of arbitrarily high order which all vanish for $X = 0$.

Since X is supposed to be $\neq R^2$, we know that $\chi^{(k)}(X)$ exists for every k and for all values of X under consideration. Hence, from (32),

(35) $$(X-Y)^{\frac{1}{2}} \chi^{(k)}(Y) = 0 \text{ for both } Y = 0 \text{ and } Y = X.$$

On writing (33) in the form

(36) $$\tau(X) = -2 \int_0^X \{d(X-Y)^{\frac{1}{2}}/dY\} \chi'(Y)\,dY/\pi; \qquad 0 \leqq X < R^2$$

and applying partial integration, the boundary condition (35) yields

(37) $$\tau(X) = 2 \int_0^X (X-Y)^{\frac{1}{2}} \chi''(Y)\,dY/\pi; \qquad 0 \leqq X < R^2.$$

---

* Cf. the definitive results of L. Tonelli, "Su un problema di Abel," *Mathematische Annalen*, Vol. 99 (1928), pp. 185-192.

Hence $\tau'(X)$ exists, viz.

$$(38) \qquad \tau'(X) = \int_0^X (X-Y)^{-\frac{1}{2}} \chi''(Y)\,dY/\pi; \qquad 0 \leqq X < R^2.$$

Since all derivatives of $\chi(Y)$ exist for $0 \leqq Y \leqq X$ and (35) holds for every k, the process which led from (33) to (38) may be repeated indefinitely, i. e. all derivatives $\tau^{(k)}(X)$ exist and

$$(39) \qquad \tau^{(k)}(X) = \int_0^X (X-Y)^{-\frac{1}{2}} \chi^{(k+1)}(Y)\,dY/\pi; \qquad 0 \leqq X < R^2.$$

Finally from (39)

$$(40) \qquad \tau^{(k)}(0) = 0.$$

Q. E. D.

We now prove (28). The even momentum

$$\int_0^R r^{2n} \rho'(r)\,dr$$

of ρ' is, according to (26),

$$= 2 \int_0^R r^{2n} \left[\int_r^R (q^2 - r^2)^{-\frac{1}{2}} q\delta(q)\,dq \right] dr,$$

i. e., by Dirichlet's rule,*

$$= \int_0^R \left[2 \int_0^q r^{2n} (q^2 - r^2)^{-\frac{1}{2}}\,dr \right] q\delta(q)\,dq$$

or (on placing $r = qp$ where q is fixed)

$$= \int_0^R \left[2 \int_0^1 (qp)^{2n} (1-p^2)^{-\frac{1}{2}}\,dp \right] q\delta(q)\,dq.$$

Hence

$$\int_0^R r^{2n} \rho'(r)\,dr = \left[\int_0^R q^{2n}\, q\delta(q)\,dq \right] \left[2 \int_0^1 p^{2n} (1-p^2)^{-\frac{1}{2}}\,dp \right],$$

where †

$$\left[2 \int_0^1 p^{2n} (1-p^2)^{-\frac{1}{2}}\,dp \right] = 2 \int_0^{\pi/2} \cos^{2n}\theta\,d\theta = \pi (2n)\,!/(n\,!^2\, 2^{2n}).$$

Accordingly,

$$\pi \int_0^R r^{2n+1} \delta(r)\,dr/(n\,!^2\, 2^{2n}) = \int_0^R r^{2n} \rho'(r)\,dr/(2n)\,!$$

or

$$\pi \int_0^R (-s^2 r^2/4)^n\, r\delta(r)\,dr/(n\,!^2) = \int_0^R (-s^2 r^2)^n \rho'(r)\,dr/(2n)\,!,$$

---

* Cf., for instance, L. Tonelli, *loc. cit.*

† Cf. in this connection G. Pólya, "Application of a theorem connected with the problem of moments," *The Messenger of Mathematics*, Vol. 55 (1926), pp. 189-192.

where s is arbitrary. This may be written, by virtue of the developments

$$J_0(sr)=\sum_{n=0}^{\infty}(-s^2r^2/4)^n/(n!^2), \qquad \cos(sr)=\sum_{n=0}^{\infty}(-s^2r^2)^n/(2n)!,$$

in the form

$$\pi\int_0^R J_0(sr)r\delta(r)\,dr=\int_0^R\cos(sr)\rho'(r)\,dr,$$

the legality of the term-by-term integration being trivial. Hence from (20)

$$(41)\qquad 2\pi\int_0^R J_0(sr)r\delta(r)\,dr=L(s;\,\rho).$$

On the other hand,*

$$\int_0^{2\pi}\exp\{i\cos(w\vartheta)\}\,d\vartheta=2\pi J_0(w),$$

i. e.

$$(42)\qquad \int_0^{2\pi}\exp\{i(u\cos\vartheta+v\sin\vartheta)r\}\,d\vartheta=2\pi J_0(rs)\quad\text{where } s=\{u^2+v^2\}^{1/2}.$$

On substituting (42) in (41) there results (28).

The continuous function $\delta(r)$ has so far been defined for $0\leqq r\leqq R$ only. It will be convenient to set

$$(43)\qquad \delta(r)=0\quad\text{for}\quad R<r<+\infty.$$

By virtue of (27) this extended function δ possesses derivatives of any order for $0<r<+\infty$.

The Laplace Transform of the Time Averages.

It is supposed that the frequencies λ_j of (1) are linearly independent. Hence if n, m, k denote arbitrary non-negative integers,

$$(44)\qquad \lim_{T=+\infty}(1/2T)\int_{-T}^{T}\Big[\sum_{j=1}^{k}r_j\cos\lambda_j(t-t_j)\Big]^n\Big[\sum_{j=1}^{k}r_j\sin\lambda_j(t-t_j)\Big]^m dt$$
$$=(1/2\pi)^k\int_0^{2\pi}\cdots\int_0^{2\pi}\Big[\sum_{j=1}^{k}r_j\cos\theta_j\Big]^n\Big[\sum_{j=1}^{k}r_j\sin\theta_j\Big]^m d\theta_1\cdots d\theta_k,$$

where $\theta_1,\cdots,\theta_k$ are k independent integration variables. This well-known identity may be verified either by complete induction or else directly and yields, according to Bohr, a simple proof for the Kronecker approximation theorem. We shall use (44) as in the paper *II* for purposes which are finer ‡

* Cf. Courant-Hilbert, *op. cit.*, p. 390.

† Cf. E. C. Titchmarsh, *The zeta-function of Riemann*, Cambridge University Press, 1930, p. 98.

‡ Cf. the introduction of the paper *II*, referred to on p. 310.

than the Kronecker theorem. In fact, we shall extend the *statistical* relation (14) to the case of the complex-valued distribution (1).

Let $f(t) = g(t) + ih(t)$ be an almost-periodic (hence * continuous and bounded) function of the real variable t, where g and h are real, and let $\{f_k(t)\}$ denote a sequence of such functions. The exponential

$$\exp i\{ug(t) + vh(t)\} \quad \text{where} \quad f = g + ih \tag{45}$$

is an almost-periodic function of t for all real and complex values of the parameters u, v, inasmuch as (45) is a uniform limit † of such functions; in fact,

$$\| g^n h^m \| \leqq \| g + ih \|^{n+m}, \tag{46}$$

where $\| q \|$ denotes the least upper bound of $| q(t) |$ in the infinite range $-\infty < t < +\infty$. Obviously

$$\lim_{k=\infty} \| \exp i\{ug_k + vh_k\} - \exp i\{ug + vh\} \| = 0 \quad \text{whenever} \lim_{k=\infty} \| f_k - f \| = 0, \tag{47}$$

where $f_k = g_k + ih_k$ and $f = g + ih$. The operator

$$\mathfrak{M}(f) = \lim_{T=+\infty} \int_{-T}^{T} f(t)\, dt/2T \tag{48}$$

is defined ‡ for every almost-periodic function f, hence for the function (45). For the time-average of this exponential we introduce the abbreviation

$$\mathfrak{L}(u, v;\, f) = \mathfrak{M}(\exp i\{ug + vh\}) \quad \text{where} \quad f = g + ih, \tag{49}$$

so that $\mathfrak{L}$ may be considered as the Laplace-Fourier transform of the time-function $f(t)$. Clearly

$$\lim_{k=\infty} \mathfrak{M}(f_k) = \mathfrak{M}(f) \quad \text{whenever} \quad \lim_{k=\infty} \| f_k - f \| = 0. \tag{50}$$

Also, for all values of the parameters u, v,

$$\mathfrak{L}(u, v;\, f) = \sum_{p=0}^{\infty} p!^{-1} \sum_{q=0}^{p} C_{pq} (iu)^{p-q} (iv)^q\, \mathfrak{M}(g^{p-q} h^q), \tag{52}$$

where $f = g + ih$ and

$$C_{pq} = p!\, (p-q)!^{-1}\, q!^{-1}. \tag{53}$$

The development (52), resulting formally from (49), is legalized by (50) and (46); in fact, $g(t)^n h(t)^m$ is § an almost-periodic function as $f(t) = g(t) + ih(t)$ is.

---

* H. Bohr, "Fastperiodische Funktionen," *Ergebnisse der Mathematik und ihre Grenzgebiete*, Vol. 1, No. 5 (1932), pp. 29-30.

† H. Bohr, *ibid.*, pp. 31-33.

‡ H. Bohr, *ibid.*, pp. 34-36.

§ H. Bohr, *ibid.*, p. 33.

Let

$$(55)\qquad z_k(t) = x_k(t) + iy_k(t) = \sum_{j=1}^{k} c_j(t)$$

denote a partial sum of (1), where

$$(56)\qquad r_j \exp i\lambda_j(t - t_j) = r_j \cos \lambda_j(t - t_j) + ir_j \sin \lambda_j(t - t_j) = c_j.$$

Then $\lim_{k=\infty} \| z_k - z \| = 0$ is assured by (2). Hence from (47), (50), (49)

$$(57)\qquad \mathfrak{L}(u, v;\ z) = \lim_{k=\infty} \mathfrak{L}(u, v;\ z_k).$$

Furthermore,

$$(58)\qquad \mathfrak{L}(u, v;\ c_j) = (1/2\pi) \int_0^{2\pi} \exp i\{(u \cos \theta + v \sin \theta) r_j\}\, d\theta.$$

In fact, (58) holds by virtue of (56) and (52) if and only if

$$(59)\qquad \mathfrak{M}([r_j \cos \lambda_j(t - t_j)]^{p-q} [r_j \sin \lambda_j(t - t_j)]^q)$$
$$= (1/2\pi) \int_0^{2\pi} [r_j \cos \theta]^{p-q} [r_j \sin \theta]^q\, d\theta \qquad (p \geqq q \geqq 0),$$

where we developed the integral $\int_0^{2\pi}$ occurring in (58) according to the powers of u and v. Since j has in (59) a fixed value it is sufficient to prove (59) for $j = 1$, and on placing in (44)

$$n = p - q, \qquad m = q, \qquad k = 1,$$

there results (59) for $j = 1$. Hence (58) holds true. Also, from (44), (55) and (56),

$$(60)\qquad \mathfrak{M}(x_k^{p-q} y_k^q)$$
$$= (1/2\pi)^k \int_0^{2\pi} \cdots \int_0^{2\pi} [\sum_{j=1}^{k} r_j \cos \theta_j]^{p-q} [\sum_{j=1}^{k} r_j \sin \theta_j]^q\, d\theta_1 \cdots d\theta_k$$

where $p \geqq q \geqq 0$.

On replacing in (52) the typical function $f(t) = g(t) + ih(t)$ by the function (55), it follows from (60) that

$$\mathfrak{L}(u, v;\ z_k) = \sum_{p=0}^{\infty} p!^{-1} \sum_{q=0}^{p} C_{pq} (iu)^{p-q} (iv)^q (1/2\pi)^k$$
$$\times \int_0^{2\pi} \cdots \int_0^{2\pi} [\sum_{j=1}^{k} r_j \cos \theta_j]^{p-q} [\sum_{j=1}^{k} r_j \sin \theta_j]^q\, d\theta_1 \cdots d\theta_k.$$

Accordingly from (53)

$$\mathfrak{L}(u, v;\ z_k) = (1/2\pi)^k \sum_{p=0}^{\infty} p!^{-1}$$
$$\times \int_0^{2\pi} \cdots \int_0^{2\pi} \{iu \sum_{j=1}^{k} r_j \cos \theta_j + iv \sum_{j=1}^{k} r_j \sin \theta_j\}^p\, d\theta_1 \cdots d\theta_k,$$

which may be written in the form

$$\mathfrak{L}(u,v;\,z_k) = (1/2\pi)^k$$
$$\times \int_0^{2\pi} \cdots \int_0^{2\pi} \sum_{p=0}^{\infty} p\,!^{-1}\{\sum_{j=1}^{k} (iur_j \cos\theta_j + ivr_j \sin\theta_j)\}^p \, d\theta_1 \cdots d\theta_k,$$

the legality of the term-by-term integration being trivial. Consequently,

$$\mathfrak{L}(u,v;\,z_k)$$
$$= (1/2\pi)^k \int_0^{2\pi} \cdots \int_0^{2\pi} \exp \sum_{j=1}^{k} (iur_j \cos\theta_j + ivr_j \sin\theta_j)\, d\theta_1 \cdots d\theta_k,$$

or

$$= (1/2\pi)^k \int_0^{2\pi} \cdots \int_0^{2\pi} \prod_{j=1}^{k} \exp(iur_j \cos\theta_j + ivr_j \sin\theta_j)\, d\theta_j,$$

i. e.

$$\mathfrak{L}(u,v;\,z_k) = \prod_{j=1}^{k} (1/2\pi) \int_0^{2\pi} \exp ir_j\{u \cos\theta_j + v \sin\theta_j\} d\theta_j.$$

Hence from (58) and (57)

$$\mathfrak{L}(u,v;\,z) = \prod_{j=1}^{\infty} \mathfrak{L}(u,v;\,c_j). \tag{61}$$

The multiplicative rule (61) is analogous to (14). The expressions $\mathfrak{L}$ are, however, time-averages whereas the integrals L represent space-integrals extended over the one-dimensional phase-space.* We shall now transform the time-averages $\mathfrak{L}$ in space-integrals Λ extended over the present phase-space which is the plane (x,y).

The Statistical Independence.

Let R denote the least upper bound of $|\,z(t)\,|$, where $z(t) = x(t) + iy(t)$ is an almost-periodic function. We do not suppose, at present, that the frequencies λ_j are linearly independent. Let Q be a rectangle in the (x,y)-plane parallel to the coördinate axes, and let $\{Q;\,T\}$ denote the set of those values t in the interval $|\,t\,| < T$ for which the point $x = x(t)$, $y = y(t)$ is within Q. In a recent paper Haviland † proves the following theorems:

(I). Every almost-periodic function $z(t)$ does possess a distribution function. In a more precise manner, there exists a monotone ‡ absolutely additive § set-function $\phi(E)$ such that

* Cf. in this connection G. D. Birkhoff, "Proof of the Ergodic Theorem," *Proceedings of the National Academy of Sciences*, Vol. 17 (1931), pp. 650-660.

† E. K. Haviland, *loc. cit.* The order of presentation of these theorems differs in his paper from that given here.

‡ J. Radon, "Theorie und Anwendungen der absolut additiven Mengenfunktionen," *Sitzungsberichte der mathematisch-naturwissenschaftlichen Klasse der Kaiserl. Akademie der Wissenschaften zu Wien*, Vol. 122 (1913), pp. 1295-1438 (more particularly p. 1303) and "Ueber lineare Funktionaltransformationen und Funktionalgleichungen," *ibid.*, Vol. 128 (1919), pp. 1083-1121.

§ J. Radon, *loc. cit.*, p. 1299.

$$\lim_{T=+\infty} \text{meas}\,\{Q;\,T\}/2T \text{ exists and } = \phi(Q),$$

provided that none of the four boundary lines of Q lies on a certain denumerable set of lines $x = x_j$, $y = y_k$. These are termed singular lines of ϕ.

(II). These lines cannot exist if * the total variation of $\phi(E)$ in Q is an absolutely continuous set-function of Q. On the other hand, there exist † almost-periodic functions $z(t)$ having actually a singular line $x = x_j$ or $y = y_k$.

(III). Since $|z(t)| \leqq R$ for every t, it is clear from (I) that $\phi(E)$ vanishes for all rectangles $E = Q$ without the circle $x^2 + y^2 \leqq R^2$. Hence ‡ the double Stieltjes integral

$$\int\int_{-\infty}^{+\infty} P(x, y)\, d\phi(E)$$

exists for every continuous point-function $P(x, y)$. In particular, all momenta

$$\int\int x^n y^m d\phi(E); \qquad (n, m = 0, 1, 2, \cdots)$$

of ϕ exist. Here and always if not otherwise indicated the integration is extended over any region containing the circle $x^2 + y^2 \leqq R^2$, e. g. over the whole (x, y)-plane.

(IV). The momenta of $\phi(E)$ are the corresponding time-momenta of $z(t) = x(t) + iy(t)$:

$$\int\int x^n y^m\, d\phi(E) = \lim_{T=+\infty} (1/2T) \int_{-T}^{T} x(t)^n\, y(t)^m\, dt,$$

where $n, m = 0, 1, 2, \cdots$.

(V). If an absolutely additive set-function $\omega(E)$ vanishes § for all rectangles without a sufficiently large circle, and if the momenta of $\omega(E)$ represent the corresponding time-momenta of $z(t)$, then ω is identical ¶ with the distribution function ϕ of $z(t)$ although it is not *presupposed* that ω be monotone.

---

* Cf. J. Radon, *loc. cit.*, pp. 1320-1322 and pp. 1093-1094.

† Cf. Bohr's example referred to above (p. 311).

‡ Cf. J. Radon, *loc. cit.*, pp. 1322-1324.

§ It may be shown that this restriction can be omitted. We do not need, however, this extension of the uniqueness theorem.

¶ This is to mean that $\omega(Q) = \phi(Q)$ holds for all those rectangles Q which are not excluded by (I). The actual value of the monotone set-function ϕ for the "singular" rectangles is undetermined and immaterial in the same sense as is the actual value of a monotone function $\rho(\xi)$ at a discontinuity point $\xi = \xi_m$. Cf. the papers of Radon and Haviland, referred to above.

These theorems of Haviland correspond to those results regarding a real-valued almost-periodic function which are proven in my first paper, referred to on p. 309, footnote †. We know that the latter results may essentially be refined if the frequencies λ_j be linearly independent. In this case we found explicit results instead of the mere existence theorem (3). We shall now extend these explicit results to complex-valued almost-periodic functions with linearly independent frequencies. This case is of first importance in the analytic theory of numbers. Even without the assumption of linear independence we have as a consequence of Haviland's results the following

LEMMA. *Let $\phi(E)$ denote the distribution function of the almost-periodic function $z(t)$. Set*

$$\Lambda(u, v;\, \omega) = \iint \exp i\{ux + vy\}\, d\omega(E) \tag{62}$$

where $\omega(E)$ is any absolutely additive set-function vanishing without a sufficiently large circle $x^2 + y^2 \leqq R^2$. Then *

$$\omega = \phi \tag{63a}$$

holds if and only if

$$\Lambda(u, v;\, \omega) = \mathfrak{L}(u, v;\, z) \tag{63b}$$

for all values of the arbitrary parameters u and v.

In fact, on placing

$$\mathrm{M}_{nm}(\omega) = \iint x^n y^m d\omega(E),$$

we have from (62) and (53)

$$\Lambda(u, v;\, \omega) = \sum_{p=0}^{\infty} p\,!^{-1} \sum_{q=0}^{p} C_{pq}(iu)^{p-q}(iv)^q\, \mathrm{M}_{p-q\,q}(\omega),$$

the legality of the term-by-term integration being trivial. On the other hand, from (52),

$$\mathfrak{L}(u, v;\, z) = \sum_{p=0}^{\infty} p\,!^{-1} \sum_{q=0}^{p} C_{pq}(iu)^{p-q}(iv)^q\, \mathfrak{M}(x^{p-q}y^q),$$

where $z(t) = x(t) + iy(t)$. On comparing the coefficients of these integral power series we see that (63b) is equivalent to

$$\mathrm{M}_{nm}(\omega) = \mathfrak{M}(x^n y^m); \qquad (n, m = 0, 1, 2, \cdots). \tag{63c}$$

Now (63c) follows from (63a) by (IV), and (63a) follows from (63c) by (V), so that (63a) is equivalent to (63c). Hence (63a) is equivalent to (63b).

* Cf. the previous footnote.

According to the Lemma thus proven we have

$$\mathfrak{L}(u, v;\, z) = \Lambda(u, v;\, \phi) \tag{64}$$

and

$$\mathfrak{L}(u, v;\, c_j) = \Lambda(u, v;\, \psi_j), \tag{65}$$

where ψ_j denotes the distribution function of the periodic function $c_j(t) = r_j \exp i\lambda_j(t - t_j)$. On substituting (64) and (65) in (61) we obtain the statistical independence relation

$$\Lambda(u, v;\, \phi) = \prod_{j=1}^{\infty} \Lambda(u, v;\, \psi_j), \tag{66}$$

which is by virtue of (1) and (56) the two-dimensional analogue of (14).

On comparing (66) with (14) and using the Abel integral equation (26) we shall now *calculate* the distribution function of (1) in terms of the one-dimensional distribution function ρ, which we know by the explicit representation (23). It would not be difficult to consider spaces with more than two dimensions. Besides, the treatment of spaces with an odd number of dimensions is simpler insofar as no Abel integral equation occurs. The occurance of this integral equation in the case of an even dimension number is related to well-known facts regarding Huyghens' Principle.*

The Distribution Function.

The total variation of a distribution function $\phi(E)$ belonging to an arbitrary almost-periodic function $z(t)$ is $= 1$. In fact, on placing both exponents n, m in (IV) equal to zero, there results

$$\iint d\phi(E) = 1. \tag{67}$$

Since $\phi(E)$ is by (I) monotone and $\geqq 0$ we conclude that

$$0 \leqq \phi(E) \leqq 1 \tag{68}$$

for every E.

Let $D(x, y)$ be a continuous point-function which is $= 0$ when $x^2 + y^2 \geqq R^2$. Then

$$\omega(E) = \iint_E D(x, y)\, dx\, dy \tag{69}$$

* Cf. Philomena Mader, "Ueber die Darstellung von Punktfunktionen im n-dimensionalen euklidischen Raum durch Ebenenintegrale," *Mathematische Zeitschrift*, Vol. 26 (1927), pp. 646-652. This paper contains also references to previous investigations. Cf. also J. Hadamard, *Le problème de Cauchy et les équations aux dérivées partielles linéaires hyperboliques*, Paris, 1932, *passim*.

is an absolutely additive set-function which vanishes for all rectangles without the circle $x^2 + y^2 \leqq R^2$. Furthermore,

$$\int\int P(x, y)\, d\omega(E) = \int\int P(x, y) D(x, y)\, dx\, dy \tag{70}$$

for any continuous point-function $P(x, y)$. For we have from (69)

$$\omega(Q_{ik}) = D(\xi_{ik}, \eta_{ik}) \mid Q_{ik} \mid, \tag{69a}$$

where (ξ_{ik}, η_{ik}) is some point in the interior or on the boundary of the rectangle Q_{ik}, and $\mid Q_{ik} \mid$ denotes the area of Q_{ik}. Accordingly,

$$\sum_i \sum_k P(\xi_{ik}, \eta_{ik}) \omega(Q_{ik}) = \sum_i \sum_k P(\xi_{ik}, \eta_{ik}) D(\xi_{ik}, \eta_{ik}) \mid Q_{ik} \mid \tag{70a}$$

for every partition of the square $\mid x \mid \leqq R$, $\mid y \mid \leqq R$ in rectangles Q_{ik}. On considering a sequence of partitions in such a way that the maximum diameter of the rectangles occurring in the n-th partition approaches zero when $\lim n = \infty$, equation (70) follows from (70a) by the integral definitions of Radon and Riemann respectively.

If the distribution function $\phi(E)$ of an almost-periodic function $z(t)$ possesses a representation (69), it is clear from (II) that the sequence of singular lines mentioned under (I) cannot exist, i. e. that

$$\lim_{T=\infty} \text{meas}\,\{Q;\ T\}/2T = \int\int_Q D(x, y)\, dx\, dy \tag{71}$$

holds for *every* Q. If the frequencies of the almost-periodic function $z(t)$ be linearly independent, its distribution function may be represented, according to Jessen, in the form (69), provided that $z(t)$ is analytic by virtue of its representation as an absolutely convergent Dirichlet series (cf. p. 310 above). The distribution function $\psi_j(E)$ belonging to the partial vibration (56) of (1) does not allow a representation (69). More than that, there does not exist a measurable function possessing over E a Lebesgue integral $= \psi_j(E)$. In fact, the very definition of a distribution function yields from (56) the relation

$$2\pi r_j \psi_j(E) = \text{length of the arc } E_j, \tag{72}$$

where E_j denotes that portion of the circle $x^2 + y^2 = r_j^2$ which is within the open rectangle E, provided that there exist such a portion; otherwise $\psi_j(E) = 0$. Now this set-function is clearly not absolutely continuous and therefore does not allow a Lebesgue representation.

From (72) and (62) we obtain by the Radon integral definition the formula

(73) $$\Lambda(u, v;\, \psi_j) = \int_0^{2\pi} \exp\{i(ur_j \cos\theta + vr_j \sin\theta)\}\; d\theta/2\pi$$

where $x = r\cos\theta$, $y = r\sin\theta$. Besides, (73) follows from (58) and (65) also. Now from (73) and (42)

(74) $$\Lambda(u, v;\, \psi_j) = J_0(r_j\{u^2 + v^2\}^{1/2}).$$

Hence, from (66),

(75) $$\Lambda(u, v;\, \phi) = \prod_{j=1}^{\infty} J_0(r_j\{u^2 + v^2\}^{1/2}).$$

On comparing (75) with (16) there results

(76) $$\Lambda(u, v;\, \phi) = L(\{u^2 + v^2\}^{1/2};\, \rho),$$

or, according to (28),

(77) $$\Lambda(u, v;\, \phi) = \int_0^R \int_0^{2\pi} r\delta(r) \exp\{i(u\cos\vartheta + v\sin\vartheta)r\}\; dr\, d\vartheta.$$

On placing

(78) $$x = r\cos\vartheta, \qquad y = r\sin\vartheta$$

and applying (70) to the point-function

$$P(x, y) = \exp\{i(ux + vy)\} = \exp\{i(u\cos\vartheta + v\sin\vartheta)r\}$$

and the absolutely additive set-function

(79) $$\omega(E) = \int\int_E \delta(\sqrt{x^2 + y^2})\,dx\,dy;\; \delta(\sqrt{x^2 + y^2}) = 0 \text{ for } x^2 + y^2 \geqq R^2,$$

we see from (62) that

(80) $$\Lambda(u, v;\, \omega) = \int\int \delta(r)\exp\{i(u\cos\vartheta + v\sin\vartheta)r\}dx\,dy;\; x^2 + y^2 = r^2.$$

Since $dx\,dy = r\,dr\,d\vartheta$, it is clear from (43) that the double integrals occurring in (77) and (80) are identical. Consequently

$$\Lambda(u, v;\, \phi) = \Lambda(u, v;\, \omega),$$

or, according to (64),

$$\mathfrak{L}(u, v;\, z) = \Lambda(u, v;\, \omega).$$

Hence from the Lemma

(81) $$\phi(E) = \omega(E)$$

(p. 324). Since ϕ is monotone by (I) it is clear from (79) that for the distribution function (81) the singular lines not excluded by (I) cannot exist and that (81) holds for *every* rectangle. On comparing (69) with (79) we see that $D(x, y) = \delta(r)$, i.e. that the distribution of (1) is of central sym-

metry. This is in accordance with the Kronecker-Weyl approximation theorem. Furthermore, from (79),

$$\delta(r) \geqq 0, \tag{82}$$

inasmuch as (81) is monotone by (I).

Accordingly, the asymptotic distribution of the values of every almost-periodic function

$$z(t) = x(t) + iy(t) = \sum_{j=1}^{\infty} r_j \exp i(t - t_j)\lambda_j; \; r_j > 0, \; R = \sum_{j=1}^{\infty} r_j < +\infty$$

with linearly independent frequencies λ_j possesses a non-negative density of probability δ which is a function of $r^2 = x^2 + y^2$ alone. This function $\delta(r)$ possesses derivatives of arbitrarily high order if $r \neq 0$ and remains continuous at the origin $r = 0$. The radial density is explicitly given by the formula

$$\pi\delta(r) = -\int_r^R (q^2 - r^2)^{-\frac{1}{2}} \rho''(q)\,dq; \; \delta(r) = 0 \textit{ for } r \geqq R, \tag{83a}$$

where *

$$\pi\rho''(r) = -\int_0^{+\infty} q \sin(rq) \prod_{j=1}^{\infty} J_0(r_j q)\,dq; \; \rho''(r) = 0 \textit{ for } r \geqq R. \tag{83b}$$

Also,

$$\pi\rho^{(2n+1)}(r) = (-1)^n \int_0^{+\infty} q^{2n}\, \Xi(q) \cos(rq)\,dq, \qquad (n = 0, 1, 2, \cdots), \tag{84a}$$

and

$$\pi\rho^{(2n)}(r) = (-1)^n \int_0^{+\infty} q^{2n-1}\, \Xi(q) \sin(rq)\,dq, \qquad (n = 1, 2, \cdots), \tag{84b}$$

where †

$$\Xi(q) = \prod_{j=1}^{\infty} J_0(r_j q) \tag{85}$$

and

$$\Xi(q) = O(q^{-m}) \textit{ when } q \to +\infty \tag{86}$$

for every fixed value of $m \geqq 0$.

The important point is that the Radon integral notion allows the treatment of "discontinuous" distributions of the type (72). The method is valid also in the case, illustrated by a geometrical investigation by Bohr and

* Cf. p. 316 and p. 315 above.

† The product (85) governs also some other statistical problems. Cf. Lord Rayleigh, "On the problem of random vibrations, and of random flights in one, two, three dimensions," *Philosophical Magazine*, Series 6, Vol. 37 (1919), pp. 321-347; R. Lüneburg, "Das Problem der Irrfahrt ohne Richtungsbeschränkung und die Randwertaufgabe der Potentialtheorie," *Mathematische Annalen*, Vol. 104 (1931), p. 700 etc.

Jessen (referred to on p. 316), where the densities are distributed along arbitrary convex curves. Applications to the ζ-function will be given later on.*

The Radial Distribution Function.

The modulus of (1) is, as (1), an almost-periodic function.† On the other hand, on replacing $z(t)$ by $|z(t)|$ we lose the linear independence of the frequences. It is nevertheless possible to calculate the distribution function of $|z(t)|$, i. e. the radial distribution function of $z(t)$. In fact, on placing

$$(87)\quad \nu(\xi)=0,\ -\infty<\xi<0;\quad \nu(\xi)=\int_0^{\xi}\eta\delta(\eta)\,d\eta/2\pi,\ 0\leqq\xi<+\infty,$$

where δ is given by (83a) and (83b), it is easy to prove that ‡

$$(88)\qquad \mathfrak{M}(|z|^n)=\int_{-\infty}^{+\infty}\xi^n\,d\nu(\xi),\qquad (n=0,1,2,\cdots).$$

Hence § $\nu(\xi)$ is the distribution function of $|z(t)|$. Thus the radial symmetry of $\phi(E)$ may be interpreted as an indication of the existence of a "mean motion" for the function $\arg z(t)$ although

$$\exp[i\arg z(t)]=z(t)/|z(t)|.$$

need not be almost-periodic.¶

We shall not use here all momentum equations (88) but only the relation

$$(89)\qquad \int_0^R r\delta(r)\,dr=2\pi,\qquad (n=0)$$

which is an obvious consequence of (67), (81), (79), (78) and (70).

---

* Cf. H. Bohr und R. Courant, "Neue Anwendungen der Theorie der diophantischen Approximationen auf die Riemannsche Zetafunktion," *Journal für Mathematik*, Vol. 144 (1914), pp. 249-274; H. Bohr und B. Jessen, "Ueber die Wertverteilung der Riemannschen Zetafunktion," *Acta Mathematica*, Vol. 54 (1930), pp. 1-35 and Vol. 58 (1932), pp. 1-55.

† This follows from the definition of the almost-periodicity inasmuch as

$$||z(t+a)|-|z(t)||\leqq|z(t+a)-z(t)|.$$

‡ The verification may be based upon the momentum identities developed in the Chapter on the Abelian integral equation.

§ Cf. *loc. cit. I* (referred to on p. 309).

¶ Cf. H. Weyl, "Sur une application de la théorie des nombres a la mécanique statistique et la théorie des perturbations," *L'Enseignement Mathématique*, Vol. 16 (1914), pp. 455-467. Cf. also F. Bernstein, "Ueber eine Anwendung der Mengenlehre auf ein aus der Theorie der säkularen Störungen herrührendes Problem," *Mathematische Annalen*, Vol. 71 (1912), pp. 417-439; and, on the other hand, H. Bohr, "Kleinere Beiträge zur Theorie der fastperiodischen Funktionen. I.," *Det Kgl. Danske Videnskabernes Selskab. Meddelelser*, Vol. 10, No. 10 (1930).

It is clear from (34), (82) and (89) that the second derivative of $\rho(\xi)$ is non-positive and not identically zero in a certain vicinity $R-\epsilon \leqq \xi \leqq R$ of the end-point $\xi = R$. It would be interesting to know if it is allowed to place $\epsilon = R$. This would mean that ρ represents, as does the Gauss curve, a so-called symmetrically convex distribution, i. e. one such that the density of probability is a non-increasing function of the distance from the origin. A detailed discussion of the curve ρ ought to be based * upon the Fourier integrals (84a), (84b).

A Property of Real Lagrangian Repartitions.

It has been pointed out in connection with (25) that the function $\rho(r)$ cannot be constant in the vicinity of points which are within the range $0 \leqq r \leqq R$. Also, the function $\rho(r)$ has derivatives of any order for all values of r. We now show that $\rho(r)$ need not be an analytic function in the range $0 \leqq r \leqq R$, even if $z(t)$ be analytic by virtue of its representation as an absolutely convergent Dirichlet series (cf. the Introduction).

Suppose that one of the partial vibrations of (1) or (5), say the first one $(j=1)$, is "overwhelming" in the sense of Lagrange: †

$$r_1 > \sum_{j=2}^{\infty} r_j. \tag{90}$$

Then the density of probability $\rho'(r)$ belonging to $x(t)$ is a positive constant in the range

$$0 \leqq r \leqq r_1 - \sum_{j=2}^{\infty} r_j (< R), \tag{91}$$

without being a constant in the whole range $0 \leqq r \leqq R$, i. e. *the repartition of* $x(t)$ *is an equipartition in the domain* (91) *but not in the whole domain of* $x(t)$. This is, in reality, a consequence of (23) but the proof is shorter if we use $\delta(r)$.

First, from (1), (2) and (90),

$$|z(t)| \geqq r_1 - \sum_{j=2}^{\infty} r_j > 0,$$

i. e.

$$|z(t)| \geqq 2r_1 - R > 0; \qquad -\infty < t < +\infty.$$

---

* Cf. M. Mathias, "Ueber positive Fourier-Integrale," *Mathematische Zeitschrift*, Vol. 16 (1923), pp. 103-125.

† Cf. H. Bohr, "Das absolute Konvergenzproblem der Dirichletschen Reihen," *Acte Mathematica*, Vol. 36 (1913), pp. 202-209; A. Wintner, "Sur l'analyse anharmonique des inégalites séculaires fournies par l'approximation de Lagrange," *Rendiconti della R. Accademia Nazionale dei Lincei*, Series 6, Vol. 11 (1930), pp. 464-467.

Consequently, the distribution function $\phi(E)$ of $z(t)$ vanishes for all those rectangles E which are within the circle

$$x^2 + y^2 < (2r_1 - R)^2.$$

Hence from (81) and (79)

$$\delta(r) = 0 \quad \text{when} \quad 0 \leqq r \leqq 2r_1 - R,$$

or, according to (26),

$$\rho'(r) \equiv \rho'(0) \tag{92}$$

when $0 \leqq r \leqq 2r_1 - R$. On the other hand, (92) cannot hold in the *whole* range $0 \leqq r \leqq R$, i. e. the second derivative of $\rho(r)$ cannot be everywhere zero. This is obvious from (34) and (89). Finally, the constant (92) is, according to (26), equal to

$$2 \int_0^R \delta(q)\, dq,$$

and, therefore, > 0 by (82) and (89).

ADDENDUM. (May 22, 1933). During the correction of the proof sheets, Jessen published in the *Zentralblatt für Mathematik und ihre Grenzgebiete*, Vol. 6 (May 10, 1933), pp. 162-163, a review of the author's paper *III*.

Jessen states that the remark in *III* regarding the example (4) is incorrect. In reality, my remark was "Diese Bedingung kann . . ." and not "Diese Bedingung muss . . ." so that Jessen's criticism is not justified.

Jessen states that although my method is a momentum method my results *loc. cit. III* are essentially the same as those of his Thesis, referred to above (p. 309). It is clear from the present paper that the analytical, viz. *explicit* methods, as developed *loc. cit. III*, yield essentially finer results than those of Jessen. Besides, Jessen does not treat the real-valued case, which was the exclusive topic of *III*, at all, and the connection between the real-valued and the complex-valued case is also not indicated by Jessen. Finally, the work of Bohr and Jessen on the zeta-function was *loc. cit III* not overlooked but exactly referred to.

THE JOHNS HOPKINS UNIVERSITY.

ON THE ADDITION OF CONVEX CURVES IN BOHR'S THEORY OF DIRICHLET SERIES.

By E. K. Haviland.

The addition of plane convex curves has been investigated by Bohr and Jessen * for the purpose of applying the results to the analytic theory of numbers. With the use of geometrical methods, they have shown that the sum of n convex curves is a closed region bounded by a single convex curve or else a closed annular region bounded by two convex curves. Following a suggestion of A. Wintner, we propose to investigate analytically the properties of the outer curve bounding the sum by the use of supporting functions, a method which will disclose the identity with respect to the outer boundary of the vectorial addition of Bohr and Jessen with the functional addition of convex regions introduced by Brunn and Minkowski.† As a consequence, we obtain explicit formulae for the radius of curvature and relations between the lengths, also the areas, of the added convex curves and the length or the area, as the case may be, of the outer boundary of their sum, areas here referring to the areas of the convex regions bounded by the curves in question.

By the addition of two point sets is understood the vector addition of each point of one set to every point of the other.‡ In this manner, the sum of any finite number of sets may be obtained by adding them step by step. Pólya § has proved that the sum of two convex *regions* M_1 and M_2 is a convex region M such that if $h(\phi)$, $h_1(\phi)$, $h_2(\phi)$ are the supporting functions of M, M_1, M_2 respectively, then $h(\phi) = h_1(\phi) + h_2(\phi)$.

In the case of adding two convex *curves*, we obtain

Theorem I. *If two convex curves are added, the outer boundary of the resulting region forms the boundary of the convex region obtained by adding the (closed) convex regions bounded by the two original curves.*

Proof. Suppose an inner point, P_1, of M_1, when added to some point, P_2, of M_2, formed a boundary point, P_3, of $M_1 + M_2$. A point P'_1 could then be found such that $P_1 + P_2$ and $P'_1 + P_2$ were collinear with the origin O,

---

* H. Bohr and B. Jessen, "Om Sandsynlighedsfordelinger ved Addition af Konvekse Kurver," *det Kongelige Danske Videnskabernes Selskabs Skrifter, Naturvidenskabelig og Mathematisk Afdelung*, Ser. 8, Vol. 12, No. 3.

† Cf. T. Bonnesen, *Les Problèmes des Isopérimètres* (Paris, 1929), Ch. V.

‡ Cf. H. Bohr and B. Jessen, *loc. cit.*, pp. 331-332. One point of the sum arises, in general, from the addition of more than one pair of points.

§ G. Pólya, "Untersuchungen über Lücken und Singularitäten von Potenzreihen," *Mathematische Zeitschrift*, Vol. 29 (1929), pp. 572-577. We make use of the definition of Minkowski's supporting functions (Stützfunktionen) given here.

but $| O - (P'_1 + P_2) | > | O - P_3 |$, where $P_3 = P_1 + P_2$. It follows that P_3 cannot be a boundary point of $M_1 + M_2$. Hence if ω_1, ω_2, Ω denote the boundaries of M_1, M_2, M respectively, every point of Ω may be formed by the addition of a point of ω_1 and a point of ω_2, which proves the theorem.

If we assign to a convex curve the supporting function of the convex region bounded by the curve, we obtain from Theorem I and the previously quoted result of Pólya

THEOREM IIa. *If the convex curves ω_1 and ω_2 are added, forming a region whose outer boundary is the convex curve Ω, and if $h_1(\phi)$, $h_2(\phi)$, $h(\phi)$ be the supporting functions assigned to ω_1, ω_2, Ω respectively, then*

$$h(\phi) = h_1(\phi) + h_2(\phi).$$

If a third curve, ω_3, be added to the region $\omega_1 + \omega_2$, we obtain by reasoning similar to that of Theorem I a region $\omega_1 + \omega_2 + \omega_3$ whose outer boundary is the convex curve forming the outer boundary of the region $\omega_3 + \Omega$. Repeating this process step by step, we are led to

THEOREM IIb. *If the convex curves $\omega_1, \cdots \omega_n$, with the supporting functions $h_1(\phi), \cdots h_n(\phi)$ respectively, be added to form a region whose outer boundary is the convex curve Ω with the supporting function $h(\phi)$, then*

$$h(\phi) = \sum_{i=1}^{n} h_i(\phi). \tag{1}$$

For the radius of curvature of Ω in the point ϕ there then follows the explicit formula

$$r(\phi) = \sum_{i=1}^{n} r_i(\phi) = \sum_{i=1}^{n} [h_i(\phi) + h_i''(\phi)],$$

where $r(\phi)$ is the radius of curvature of Ω and $r_i(\phi)$ that of ω_i, $i = 1, \cdots n$, inasmuch as the radius of curvature is known to be $h(\phi) + h''(\phi)$ if h has continuous second derivatives.

If L denote the length of a convex curve Ω with supporting function $h(\phi)$ and if A denotes the area of the convex region bounded by the curve, then it is known that L and A can be expressed by the formulae:

$$L = \int_0^{2\pi} h(\phi)\,d\phi, \qquad A = \tfrac{1}{2}\int_0^{2\pi} [h^2(\phi) - (h'(\phi))^2]\,d\phi.$$

If we substitute the value of $h(\phi)$ given by (1) in the former equation, we obtain

$$L = \int_0^{2\pi} [\sum_{i=1}^{n} h_i(\phi)]\,d\phi = \sum_{i=1}^{n} \int_0^{2\pi} h_i(\phi)\,d\phi = \sum_{i=1}^{w} L_i,$$

where L_i is the length of the curve ω_i.

Making a similar substitution in the expression for A, we obtain

$$A = \tfrac{1}{2}\int_0^{2\pi}[(\sum_{i=1}^{n} h_i(\phi))^2 - (\sum_{i=1}^{n} h'_i(\phi))^2]d\phi$$

$$= \tfrac{1}{2}\int_0^{2\pi}\{\sum_{i=1}^{n} [h_i^2(\phi) - (h'_i(\phi))^2]\}d\phi$$

$$+ \tfrac{1}{2}\int_0^{2\pi}\{\sum_{i=1}^{n}\sum_{j=1}^{n}{}' [h_i(\phi)h_j(\phi) - h'_i(\phi)h'_j(\phi)]\}d\phi$$

$$= \sum_{i=1}^{n} \tfrac{1}{2}\int_0^{2\pi}[h_i^2(\phi) - (h'_i(\phi))^2]d\phi$$

$$+ \tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}{}'\int_0^{2\pi}[h_i(\phi)h_j(\phi) - h'_i(\phi)h'_j(\phi)]d\phi$$

$$(2) \qquad = \sum_{i=1}^{n} A_i + \sum_{i=1}^{n}\sum_{j=1}^{n}{}' M_{ij},$$

where the primes in the double summation indicate that those terms for which $i = j$ are omitted and $M_{ij} = M_{ji}$ is the so-called mixed area of Minkowski.*

Since, according to Minkowski,*

$$M_{ij} \geqq \sqrt{(A_iA_j)}$$

we obtain

$$A \geqq \sum_{i=1}^{n} A_i + \sum_{i=1}^{n}\sum_{j=1}^{n}{}' \sqrt{(A_iA_j)}, \text{ i.e., } A \geqq (\sum_{i=1}^{n} \sqrt{A_i})^2,$$

so

$$\sqrt{A} \geqq \sum_{i=1}^{n} \sqrt{A_i}.$$

We summarize these results in

THEOREM III. *If the convex curves* ω_i *of length* L_i *bounding regions of area* A_i, $i = 1, \cdots n$, *are added to form a region whose outer boundary is the convex curve* Ω *of length* L *bounding a region of area* A, *then the lengths of the curves and the areas of the regions they enclose are subject to the relations*

$$L = \sum_{i=1}^{n} L_i$$

and

$$\sqrt{A} \geqq \sum_{i=1}^{n} \sqrt{A_i}.$$

As the M_{ij} are always positive, it follows from (2) that

$$A > \sum_{i=1}^{n} A_i,$$

a result obtained by Bohr † some years ago by other methods.

THE JOHNS HOPKINS UNIVERSITY.

---

* Cf., for example, W. Blaschke, *Kreis und Kugel* (Leipzig, 1916), pp. 106-107; also the reference to T. Bonnesen made in Note (2).

† H. Bohr, "Om Addition af Uendelig Mange Konvekse Kurver," *Oversigt over det Kongelige Danske Videnskabernes Selskabs* (1913), pp. 364-365.

ON THE STABLE DISTRIBUTION LAWS.

By Aurel Wintner.

A real-valued function $\sigma(x)$ defined for $-\infty < x < +\infty$ is termed a distribution function if it satisfies the following conditions:

$$(1)\quad \sigma(x) \leqq \sigma(x+h) \text{ where } h>0; \qquad \sigma(x+0)+\sigma(x-0)=2\sigma(x);$$
$$\sigma(-\infty)=0, \qquad \sigma(+\infty)=\int_{-\infty}^{+\infty} d\sigma(x)=1.$$

It is clear that the Fourier-Laplace transform

$$(2)\qquad L(t;\sigma)=\int_{-\infty}^{+\infty}\exp(itx)\,d\sigma(x), \qquad -\infty<t<+\infty$$

exists and is a continuous * function of t. The existence of the momenta

$$(3)\qquad \mu_n=\int_{-\infty}^{+\infty} x^n d\sigma(x)$$

for $n>0$ is not supposed.

The distribution function $\rho(x)$ is said to be the statistical sum of the distribution functions $\sigma(x)$ and $\tau(x)$ if

$$(4)\qquad L(t;\rho)=L(t;\sigma)L(t;\tau)$$

holds for all values of the real parameter t. The reason for this terminology is the fact † that if σ and τ are the distribution functions of two statistically independent events x_1, x_2, then the distribution function ρ of the event x_1+x_2 satisfies the multiplicative relation (4), and conversely.

To every distribution function $\sigma(x)$ there belongs a sheaf of distribution functions

$$(5)\qquad \sigma_p(x)=\sigma(x/p) \qquad (\sigma_1=\sigma)$$

where p is an arbitrary positive number. The distribution functions belonging to the same sheaf differ from each other only in the degree of their scattering or precision. Correspondingly, if the so-called dispersion of $\sigma=\sigma_1$, viz. the integral μ_2 defined by (3), be finite, then the dispersion of σ_p is, according to (3) and (5), simply $p^2\mu_2$.

We shall restrict ourselves to distribution functions σ having a dispersion

$$(6)\qquad \mu_2=\int_{-\infty}^{+\infty} x^2 d\sigma(x) < +\infty.$$

* Cf. the footnote on the next page $(n=0)$.

† Cf., for instance, R. Deltheil, *Erreurs et moindres carrés*, Paris, 1930, p. 31.

The inequality of Schwarz then assures the absolute convergence of the integral μ_1. Since $\mu_0 = 1$ by (1) and (3) we have

$$\int_{-\infty}^{+\infty} (\xi - x\eta)^2 \, d\sigma(x) = \xi^2 - 2\mu_1\xi\eta + \mu_2\eta^2. \tag{7}$$

It is clear from (6) that (2) possesses first and second derivatives, viz.

$$L'(t;\,\sigma) = i\int_{-\infty}^{+\infty} x \exp(itx)\,d\sigma(x),$$

$$L''(t;\,\sigma) = -\int_{-\infty}^{+\infty} x^2 \exp(itx)\,d\sigma(x).$$

Furthermore, the second derivative is everywhere continuous.* Finally,

$$L(0;\,\sigma) = \mu_0 = 1, \qquad L'(0;\,\sigma) = i\mu_1, \qquad L''(0;\,\sigma) = -\mu_2. \tag{8}$$

The distribution function σ generating the sheaf (5) is said to be stable if the statistical sum of two distribution functions belonging to the sheaf is contained in this sheaf so that

$$L(t;\,\sigma_c) = L(t;\,\sigma_a)L(t;\,\sigma_b) \tag{9}$$

by virtue of the definitions (4) and (5). For instance, the Gaussian distribution

$$\sigma(x) = (2\pi\mu_2{}^2)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp(-y^2/2\mu_2{}^2)\,dy \qquad (\mu_2 \neq 0) \tag{10}$$

is a stable distribution.† Cauchy's investigations regarding the stability problem have been further developed by P. Lévy, and, in another direction, by Pólya.‡ On considering in (9) the positive numbers a, b, c not as variable

---

* In fact, if $n \geqq 0$ and

$$\int_{-\infty}^{+\infty} |\, x\,|^n \, d\sigma(x) < +\infty \tag{I}$$

then

$$\int_{-\infty}^{+\infty} x^n \exp(itx)\,d\sigma(x); \qquad -\infty < t < +\infty \tag{II}$$

represents a continuous function of t although σ may be discontinuous. Since from (I)

$$\left|\int_{\pm R}^{\pm\infty} x^n \exp(itx)\,d\sigma(x)\right| \leqq \pm\int_{\pm R}^{\pm\infty} |\,x\,|^n\,d\sigma(x) < \epsilon$$

when R is larger than a positive number depending upon ϵ but independent of t, it is sufficient to prove that

$$\int_{-R}^{R} x^n \exp(itx)\,d\sigma(x) \tag{III}$$

is a continuous function of t for every fixed value of R. Now the continuity of (III) is obvious inasmuch as $\exp(itx)$ is uniformly continuous in the rectangle $|\,x\,| \leqq R$, $|\,t\,| \leqq T$ where T is arbitrarily large but fixed.

† Cf. R. Deltheil, *op. cit.*, p. 44.

‡ G. Pólya, "Herleitung des Gaussschen Fehlergesetzes aus einer Funktionalgleichung," *Mathematische Zeitschrift*, Vol. 18 (1923), pp. 96-108.

parameters but as certain constants, Pólya proves that in this sense the Gaussian distribution is the only stable distribution satisfying (6). Pólya supposes, however, that $\sigma(x)$ possesses, up to a set of measure zero, a derivative which is bounded and integrable in the sense of Riemann in every finite range $|x| \leqq R$. Consequently it is postulated that σ be everywhere continuous, and even absolutely continuous.

This hypothesis regarding a *density* of probability is somewhat artificial inasmuch as it does not allow a direct statistical interpretation. Correspondingly, not every solution of the problem satisfies this hypothesis. For on placing

$$(11) \quad \delta(x) = 0, \; -\infty < x < 0; \quad \delta(x) = 1, \; 0 < x < +\infty; \quad \delta(0) = \tfrac{1}{2}$$

conditions (1) and (6) are clearly satisfied by $\sigma = \delta$. Furthermore, from (2) and (5),

$$L(t;\, \delta) = 1, \qquad \delta_p(x) = \delta(x),$$

so that (9) is an identity in a, b, c when $\sigma = \delta$. Thus the distribution function (11), which is the mathematical substitute for Dirac's corresponding notion, satisfies all requirements although it is discontinuous.

It is the object of the present note to point out the fact that the problem does not possess any *further* solution, i. e. that *if a distribution of finite dispersion* ($\geqq 0$) *be stable then it is either the Gaussian or else the Dirac distribution.* The stability of a distribution has to be understood as mentioned above. Otherwise, also if the dispersion be infinite, there exist infinitely many *analytic* distributions which are stable but non-Gaussian.*

First, from (2) and (5),

$$(12) \qquad L(t;\, \sigma_p) = L(pt;\, \sigma)$$

so that according to (8)

$$(13) \qquad L(0;\, \sigma_p) = 1, \quad L'(0;\, \sigma_p) = ip\mu_1, \quad L''(0;\, \sigma_p) = -p^2\mu_2.$$

On differentiating the relation (9) twice with respect to t and placing $t = 0$ it follows from (13) that

$$(14) \qquad c\,\mu_1 = a\,\mu_1 + b\,\mu_1,$$

$$(15) \qquad c^2\mu_2 = a^2\mu_2 + 2ab\mu_1^2 + b^2\mu_2.$$

Suppose that $\mu_1 \neq 0$. Then (14) yields $c = a + b$. Hence $\mu_2 = \mu_1^2$ by virtue of (15). Thus the expression (7) takes the form $(\xi - \mu_1\eta)^2$ and vanishes, therefore, at $\eta = \xi/\mu_1$, i. e.

* Cf. G. Pólya, *loc. cit.*, pp. 104-105.

$$\int_{-\infty}^{+\infty}(1-x/\mu_1)^2\,d\sigma(x)=0.$$

This is possible only * if the monotone function $\sigma(x)$ is constant in the vicinity of every point x for which $(1-x/\mu_1)^2>0$, i. e. in the vicinity of every point $x \neq \mu_1$. Hence $\sigma(x)$ is, according to (1) and (11), identical with $\delta(x-\mu_1)$. Consequently

$$L(t;\,\sigma_p)=\int_{-\infty}^{+\infty}\exp(it\{x+\mu_1\})\,d\delta(x)=\exp(it\mu_1) \tag{16}$$

by virtue of (2), (12) and (11). On substituting (16) in (9) there results

$$\exp(it\mu_1)=\exp(2it\mu_1)$$

where $\mu_1 \neq 0$ by supposition. This is a contradiction. Consequently $\mu_1=0$.

Hence (15) takes the form

$$c^2\mu_2=(a^2+b^2)\mu_2. \tag{17}$$

There are now two cases possible according as the dispersion μ_2 is or is not zero.

If $\mu_2=0$ then the monotone function $\sigma(x)$ is, by virtue † of (6), constant in the vicinity of every point $x \neq 0$. Hence $\sigma(x)$ is, according to (1), identical with the Dirac distribution function (11).

If $\mu_2 \neq 0$ then (17) may be written in the form $c^2=a^2+b^2$, and it follows with an application of Thiele's semi-invariants ‡ exactly as in Pólya's case of Riemann integrals that §

$$L(t;\,\sigma)\equiv\int_{-\infty}^{+\infty}\exp(itx)\,d\sigma(x)=\exp(-t^2\mu_2/2). \tag{18}$$

According to the uniqueness theorem ¶ of Fourier-Stieltjes integrals there cannot exist more than one distribution function σ satisfying (18). On the other hand, (18) is known ‖ to be satisfied by (10). Hence, if $\mu_2 \neq 0$, the distribution is the Gaussian one.

Thus the Dirac distribution appears in our problem not only as a Gaussian distribution of infinitely high precision ($\mu_2=0$) but also as the *only possible* limit when $\lim \mu_2=0$. For the curves themselves is this clear from (10) and (11) directly.

---

* This readily follows from the Hilfssatz 2 of O. Perron, *Die Lehre von den Kettenbrüchen*, Leipzig und Berlin, 1913, p. 368.

† Cf. the previous footnote.

‡ T. N. Thiele, *Forelaesninger over almindelig Iagttagelseslaere*, Copenhagen, 1889, pp. 16-38. Cf. also F. Hausdorff, "Beiträge zur Wahrscheinlichkeitsrechnung," *Berichte über die Verhandlungen der Königl. Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-physikalische Klasse*, Vol. 53 (1901), pp. 152-178.

§ G. Pólya, *loc. cit.*, pp. 101-102.

¶ Cf. R. Deltheil, *op. cit.*, pp. 27-29.

‖ *Ibid.*, pp. 44-45.

In this connection it is natural to ask what is the corresponding limiting distribution when $\lim \mu_2 = +\infty$. The formula (10) yields the function which is identically zero, and therefore not a distribution function inasmuch as the last condition (1) is not fulfilled. The question has, however, a meaning if we consider *angular* variables by reducing the Gaussian distribution, in the sense of Weyl,* modulo one.

First, the density of probability belonging to (10) is

$$(2\pi\mu_2{}^2)^{-1/2}\exp(-x^2/2\mu_2{}^2).$$

In order to reduce this modulo one we have to collect the probability densities of all those events x whose difference is an integer. There results thus the periodic density

$$(2\pi\mu_2{}^2)^{-1/2}\sum_{n=-\infty}^{+\infty}\exp(-(x+n)^2/2\mu_2{}^2).$$

This convergent series represents,† in Schwarz' notations, the theta-function $\vartheta_3(x \mid \pi i/2\mu_2{}^2)$. On the other hand,‡

$$\vartheta_3(x \mid \tau) = \sum_{n=-\infty}^{+\infty} h^{n^2}\cos 2\pi nx$$

where §

$$h = \exp(\tau\pi i).$$

The Gaussian density of probability of an angular variable is therefore

$$\vartheta_3(x \mid \pi i/2\mu_2{}^2) = \sum_{n=-\infty}^{+\infty}\exp(-n^2\pi^2/2\mu_2{}^2)\cos 2\pi nx.$$

The orthogonal trigonometric polynomials belonging to this periodic density in the same sense as the Hermite-Bruns polynomials belong to the Gaussian density have been determined by Szegö.¶ It is clear from the last Fourier series for ϑ_3 that the limit of the periodic density when $\lim \mu_2 = +\infty$ is

$$\vartheta_3(x \mid 0) = 1; \qquad 0 \leqq x < 1$$

due to the uniform convergence in the vicinity of $h = +0$. Hence the equipartition of Weyl ‖ may be interpreted as a *Gaussian* angular distribution of infinitely low precision $(1/\mu_2 = 0)$.

THE JOHNS HOPKINS UNIVERSITY.

* H. Weyl, "Ueber die Gleichverteilung der Zahlen mod. Eins," *Mathematische Annalen*, Vol. 77 (1916), p. 313.

† H. A. Schwarz, *Formeln und Lehrsätze zum Gebrauche der elliptischen Funktionen*, Göttingen, 1885, p. 46, formula (4).

‡ *Ibid.*, p. 41, formula (9).

§ *Ibid.*, p. 40.

¶ G. Szegö, "Ein Beitrage zur Theorie der Thetafunktionen," *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, (1926), pp. 242-252.

‖ *Loc. cit.*

THE COOLING PROBLEM FOR SPHERICAL REGIONS.

By W. M. Rust, Jr.

PART ONE. Introduction.

In a former paper * the solution of the cooling problem for several media for the case in which the heat flows in parallel lines was shown to be reducible to the solution of a set of Volterra integral equations of the second sort. The purpose of this paper is to indicate the application of this method to the problem where the flow is along radii of a sphere. The arguments which are repetitions of those in the former paper will not be given, but reference will be given to that paper.

PART TWO. The Problem.

A quantity of one material is heated and placed inside a spherical shell of another material, as a casting in its mould. The problem is: given the initial temperature of the two regions and the temperature of the outer boundary, find the temperature at any point at any time.

For convenience we take very simple conditions. We take the two surfaces of the spherical shell to be concentric. We take the initial temperature in the two regions to be constant, but not in general the same in the two regions. We take the temperature at the outer surface constant, at any instant, over the entire surface. With these conditions the temperature is symmetric about the center of the spheres and, at any instant, is constant over any sphere with that center. Thus only one spatial coördinate, the distance from the center, is involved.

As before, the conductivity of the material in the inner region is K_1 and in the outer region is K_2. The quantities a^2 and b^2 are positive constants equal to the ratios of the conductivity to the product of the specific heat by the density, for the inner and outer regions respectively. Also r is the distance from the center and t is the time after an initial time.

We take the radius of the inner sphere to be m and of the outer sphere to be l.

At the interior points the temperature in the inner region, $u_1(r, t)$, and the temperature in the outer region, $u_2(r, t)$, satisfy the partial differential equations †

---

* "Integral equations and the cooling problem for several media," *American Journal of Mathematics*, Vol. 54 (1932), p. 190.

† Carslaw, *Conduction of Heat*, page 11.

$$(2.1)\qquad \begin{cases} a^2r^2\,\partial u_1/\partial t = \partial(r^2\,\partial u_1/\partial r)/\partial r, & 0 < r < m,\ t > 0 \\ b^2r^2\,\partial u_2/\partial t = \partial(r^2\,\partial u_2/\partial r)/\partial r, & m < r < l,\ t > 0 \end{cases}$$

respectively.

If the temperature in the inner region is initially u_1, a constant, and in the outer region u_2, a constant, we have

$$(2.2)\qquad \begin{cases} \underset{t=0+}{\text{Limit}}\ u_1(r,t) = u_1 \quad \text{for} \quad 0 < r < m \\ \underset{t=0+}{\text{Limit}}\ u_2(r,t) = u_2 \quad \text{for} \quad m < r < l. \end{cases}$$

At the outer boundary the temperature is taken to be a known *absolutely continuous* function of the time, say $f(t)$. We have then

$$(2.3)\qquad \underset{r=l-0}{\text{Limit}}\ u_2(r,t) = f(t) \quad \text{for} \quad t > 0.$$

At the separating boundary we have two conditions,* first, the temperature is continuous in r across the boundary, that is,

$$(2.4)\qquad \underset{r=m-0}{\text{Limit}}\ u_1(r,t) = \underset{r=m+0}{\text{Limit}}\ u_2(r,t) \quad \text{for} \quad t > 0$$

and, second, the partial derivatives with respect to r satisfy the equation

$$(2.5)\qquad \underset{r=m-0}{\text{Limit}}\ K_1\partial u_1(r,t)/\partial r = \underset{r=m+0}{\text{Limit}}\ K_2\partial u_2(r,t)/\partial r \quad \text{for} \quad t > 0.$$

In almost the same manner used in the preceding paper for parallel flow, we establish the following Uniqueness Theorem.

UNIQUENESS THEOREM B. *There can not be more than one solution of the problem as given by equations* (2.1) *subject to the conditions* (2.2) *to* (2.5) *which is bounded everywhere (including* $t = 0$*), is continuous for* $t > 0$*, except at the boundaries, and has first derivatives with respect to each of the variables, which are continuous for* $t > 0$*, except at the boundaries, and the derivative with respect to* r *satisfies the conditions*

$$\underset{r=r_0\pm 0}{\text{Limit}} \int_{t_1}^{t_2} |\, \partial u_i(r,t)/\partial r \,|\, dt = \int_{t_1}^{t_2} \underset{r=r_0\pm 0}{\text{Limit}}\, |\, \partial u_i(r,t)/\partial r \,|\, dt, \quad \text{for} \quad t_1, t_2 > 0$$

where r_0 *is the* r *coördinate of any boundary and*

$$\underset{r=0+}{\text{Limit}} \int_{t_1}^{t_2} r^2 \,|\, \partial u_1(r,t)/\partial r \,|\, dt = 0.$$

The only difference of importance between this proof and the former one is that we take

$$J_\epsilon(t) = \frac{K_2b^2}{2}\int_{m+\epsilon}^{l-\epsilon} \{rV_2(r,t)\}^2\,dr + \frac{K_1a^2}{2}\int_{\epsilon}^{m-\epsilon} \{rV_1(r,t)\}^2\,dr$$

* Carslaw, *loc. cit.*, page 12.

where, as before, $V_t(r, t)$ is the difference between two solutions and is shown to be identically zero for $t > 0$.

We can now show that a solution of the problem satisfying the conditions of the Uniqueness Theorem B is given by

$$(2.6)\quad u_1(r,t) = u_1 + \int_0^t \{e^{-a^2(r-m)^2/4(t-t')} - e^{-a^2(r+m)^2/4(t-t')}\} r^{-1}(t-t')^{-1/2}\psi_1(t')\,dt'$$

for $0 < r < m$ and $t > 0$ and

$$(2.7)\quad u_2(r,t) = u_2 + \int_0^t e^{-b^2(r-m)^2/4(t-t')}\, r^{-1}(t-t')^{-1/2}\psi_2(t')\,dt' + \int_0^t e^{-b^2(r-l)^2/4(t-t')}\, r^{-1}(t-t')^{-1/2}\psi_3(t')\,dt'$$

for $m < r < l$ and $t > 0$, where $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$ are summable functions satisfying the following integral equations nearly everywhere.

$$(2.8)\quad \int_0^t (t-t')^{-1/2}\psi_3(t')\,dt' = f_2(t) - \int_0^t (t-t')^{-1/2} e^{-\beta^2/(t-t')}\psi_2(t')\,dt'$$

$$\begin{aligned}(2.9)\quad \psi_2(t) = &- c\psi_1(t) + c\alpha\pi^{-1/2}\int_0^t (t-t')^{-3/2} e^{-\alpha^2/(t-t')}\psi_1(t')\,dt' \\ &+ \beta\pi^{-1/2}\int_0^t (t-t')^{-3/2} e^{-\beta^2/(t-t')}\psi_3(t')\,dt' \\ &+ c\alpha^{-1}\pi^{-1/2}\int_0^t (t-t')^{-1/2}\psi_1(t')\,dt' \\ &- b^{-1}m^{-1}\pi^{-1/2}\int_0^t (t-t')^{-1/2}\psi_2(t')\,dt' \\ &- c\alpha^{-1}\pi^{-1/2}\int_0^t (t-t')^{-1/2} e^{-\alpha^2/(t-t')}\psi_1(t')\,dt' \\ &- b^{-1}m^{-1}\pi^{-1/2}\int_0^t (t-t')^{-1/2} e^{-\beta^2/(t-t')}\psi_3(t')\,dt'\end{aligned}$$

$$\begin{aligned}(2.10)\quad &\int_0^t \{(t-t')^{-1/2} e^{-\alpha^2/(t-t')} + (t-t')^{-1/2}\}\psi_1(t')\,dt' \\ &= f_1 + \int_0^t (t-t')^{-1/2}\psi_2(t')\,dt' \\ &\quad + \int_0^t (t-t')^{-1/2} e^{-\beta^2/(t-t')}\psi_3(t')\,dt'.\end{aligned}$$

The integrals are Lebesgue integrals and the following abbreviations are used:

$$\begin{aligned} \alpha &= am \\ \beta &= b(l-m)/2 \\ f_1 &= (u_2-u_1)m \\ f_2(t) &= \{f(t)-u_2\}l \\ c &= (K_1a)/(K_2b) > 0. \end{aligned}$$

By actual differentiation

$$e^{-a^2(r-r')^2/4(t-t')}\, r^{-1}(t-t')^{-\frac{1}{2}}$$

when considered as a function of r and t is a solution of the first equation (2.1) for any $r \neq r'$ and any $t' < t$. If we replace the a^2 in the exponent by b^2 we have a solution of the second equation (2.1). We see that

$$\operatorname*{Limit}_{t=t'+0} e^{-a^2(r-r')^2/4(t-t')}\, r^{-1}(t-t')^{-\frac{1}{2}} = 0$$

if $r \neq r'$ and $r \neq 0$ and so, since the equations (2.1) are linear, linear combinations such as (2.6) and (2.7) are solutions of (2.1). The integrals converge, except for $r=0$, m or l, for any summable functions $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$ since the other factors in the integrands are bounded except for these values of r. Each integrand is thus summable.

Since each integrand is summable, each integral approaches zero with t, except for the excluded values of r, hence the initial condition (2.2) is satisfied.

We now show that any set of *summable* functions satisfying the equations (2.8) to (2.10) must be of the form

$$\psi_i(t) = A_i t^{-\frac{1}{2}} + \int_0^t s_i(t')(t-t')^{-\frac{1}{2}}\,dt' \tag{2.11}$$

where A_i is a constant and $s_i(t)$ is a summable function.

Equation (2.8) has the form

$$\int_0^t (t-t')^{-\frac{1}{2}} \psi_3(t')\,dt' = g_3(t) \tag{A}$$

where

$$g_3(t) = f_2(t) - \int_0^t (t-t')^{-\frac{1}{2}} e^{-\beta^2/(t-t')}\, \psi_2(t')\,dt'.$$

The function $(t-t')^{-\frac{1}{2}} e^{-\beta^2/(t-t')}$ is bounded and is absolutely continuous in t, uniformly for all t' and so by Lemma II, page 193 of the former paper, $g_3(t)$ is absolutely continuous if $\psi_2(t)$ is summable. Hence the solution of (A) is *

$$\psi_3(t) = g_3(0)t^{-\frac{1}{2}} + \int_0^t g'_3(t')(t-t')^{-\frac{1}{2}}\,dt'$$

* Volterra, *Leçons sur les Équations Intégrales*, page 37.

where $g'_3(t)$ equals the derivative of $g_3(t)$ wherever that derivative exists, that is, nearly everywhere. Since $g'_3(t)$ is summable $\psi_3(t)$ has the required form.

Multiplying (2.9) by $(t''-t)^{-1/2}$ and integrating from $t=0$ to $t=t''$ gives

$$\text{(B)} \quad \int_0^{t''} \psi_2(t)(t''-t)^{-1/2}\,dt + c\int_0^{t''} \psi_1(t)(t''-t)^{-1/2}\,dt$$
$$= \int_0^{t''} (t''-t)^{-1/2}\,dt \left\{ c\alpha\pi^{-1/2} \int_0^t (t-t')^{-3/2} e^{-\alpha^2/(t-t')} \psi_1(t')\,dt' \right.$$
$$+ \beta\pi^{-1/2} \int_0^t (t-t')^{-3/2} e^{-\beta^2/(t-t')} \psi_3(t')\,dt'$$
$$+ c\alpha^{-1}\pi^{-1/2} \int_0^t (t-t')^{-1/2}(1-e^{-\alpha^2/(t-t')})\psi_1(t')\,dt'$$
$$- b^{-1}m^{-1}\pi^{-1/2} \int_0^t (t-t')^{-1/2} \psi_2(t')\,dt'$$
$$\left. - b^{-1}m^{-1}\pi^{-1/2} \int_0^t (t-t')^{-1/2} e^{-\beta^2/(t-t')} \psi_3(t')\,dt' \right\}.$$

A change of order of integration in each term of the right hand member will show that member is absolutely continuous if $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$ are summable.

Equation (2.10) has the form

$$\text{(C)} \quad \int_0^t (t-t')^{-1/2} \psi_1(t')\,dt' - \int_0^t (t-t')^{-1/2} \psi_2(t')\,dt'$$
$$= f_1 + \int_0^t (t-t')^{-1/2} e^{-\beta^2/(t-t')} \psi_3(t')\,dt'$$
$$- \int_0^t (t-t')^{-1/2} e^{-\alpha^2/(t-t')} \psi_1(t')\,dt'.$$

As in (A) each term of the right hand member is absolutely continuous if $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$ are summable.

Since $c \neq -1$ we can solve (B) and (C) as a pair of linear algebraic equations for

$$\int_0^t (t-t')^{-1/2} \psi_1(t')\,dt' \quad \text{and} \quad \int_0^t (t-t')^{-1/2} \psi_2(t')\,dt'.$$

In each case the solution is an absolutely continuous function and, as in (A), the functions $\psi_1(t)$ and $\psi_2(t)$ have the required form.

If we substitute these expressions into (2.6) and (2.7) we have

$$u_1(r,t) = u_1$$
$$+ A_1 \int_0^t \{e^{-a^2(r-m)^2/4(t-t')} - e^{-a^2(r+m)^2/4(t-t')}\} r^{-1}(t-t')^{-1/2}(t')^{-1/2}dt'$$

$$+\int_0^t \{e^{-a^2(r-m)^2/4(t-t')} - e^{-a^2(r+m)^2/4(t-t')}\}$$
$$\times r^{-1}(t-t')^{-\frac{1}{2}}\,dt' \int_0^{t'} s_1(t'')(t'-t'')^{-\frac{1}{2}}\,dt''$$

and

$$u_2(r,t) = u_2 + A_2 \int_0^t r^{-1}(t-t')^{-\frac{1}{2}} e^{-b^2(r-m)^2/4(t-t')}(t')^{-\frac{1}{2}}\,dt'$$
$$+\int_0^t r^{-1}(t-t')^{-\frac{1}{2}} e^{-b^2(r-m)^2/4(t-t')}\,dt' \int_0^{t'} s_2(t'')(t'-t'')^{-\frac{1}{2}}\,dt''$$
$$+A_3 \int_0^t r^{-1}(t-t')^{-\frac{1}{2}} e^{-b^2(r-l)^2/4(t-t')}(t')^{-\frac{1}{2}}\,dt'$$
$$+\int_0^t r^{-1}(t-t')^{-\frac{1}{2}} e^{-b^2(r-l)^2/4(t-t')}\,dt' \int_0^{t'} s_3(t'')(t'-t'')^{-\frac{1}{2}}\,dt''.$$

A change of order of integration in each term involving a double integral shows immediately that $u_1(r,t)$ and $u_2(r,t)$ are bounded except possibly at $r=0$. To show that $u_1(r,t)$ is bounded near $r=0$ we need to show that

$$\int_0^t (e^{-a^2(r-m)^2/4(t-t')} - e^{-a^2(r+m)^2/4(t-t')})\, r^{-1}(t-t')^{-\frac{1}{2}}(t')^{-\frac{1}{2}}\,dt'$$

is bounded.

By the definition of derivative and the fact that the derivative exists in this case, we have

$$\operatorname*{Limit}_{r=0} (e^{-a^2(r-m)^2/4(t-t')} - e^{-a^2(r+m)^2/4(t-t')})(2r)^{-1} = \left[\frac{d}{dx}(e^{-a^2x^2/4(t-t')})\right]_{x=m}$$

which is finite so that for r small enough the first two factors are bounded and so the integrand is less than $K(t-t')^{-\frac{1}{2}}(t')^{-\frac{1}{2}}$ the integral of which is bounded.

For r different from zero, say greater than $m/2$, we have

$$|\, r^{-1}(t-t')^{-\frac{1}{2}} e^{-b^2(r-l)^2/4(t-t')}\psi_3(t')\,| \leqq 2m^{-1}\,|\,\psi_3(t')\,|\,(t-t')^{-\frac{1}{2}}$$

which is summable, since $\psi_3(t)$ has the form (2.11) so that, for all $t>0$,

$$\operatorname*{Limit}_{r=l-0} \int_0^t r^{-1}(t-t')^{-\frac{1}{2}} e^{-b^2(r-l)^2/4(t-t')}\psi_3(t')\,dt'$$
$$= \int_0^t \operatorname*{Limit}_{r=l-0} \{r^{-1}(t-t')^{-\frac{1}{2}} e^{-b^2(r-l)^2/4(t-t')}\psi_3(t')\}\,dt'$$
$$= l^{-1}\int_0^t \psi_3(t')(t-t')^{-\frac{1}{2}}\,dt'$$

and so, for all $t>0$, we have

$$\text{(2.12)}\qquad \operatorname*{Limit}_{r=l-0} u_2(r,t) = u_2 + l^{-1}\int_0^t (t-t')^{-\frac{1}{2}} e^{-\beta^2/(t-t')}\psi_2(t')\,dt'$$
$$+ l^{-1}\int_0^t (t-t')^{-\frac{1}{2}}\psi_3(t')\,dt'.$$

But in virtue of equation (2. 8) this expression is equal to $f(t)$ for almost all $t > 0$. However by putting in the values for $\psi_2(t')$ and $\psi_3(t')$ given by (2. 11) we can see that $\operatorname*{Limit}_{r=l-0} u_2(r, t)$ is continuous as is $f(t)$, by hypothesis, and so the equality holds everywhere * and the boundary condition (2. 3) is satisfied for all $t > 0$.

In a similar manner we can show that $\operatorname*{Limit}_{r=m-0} u_1(r, t)$ and $\operatorname*{Limit}_{r=m+0} u_2(r, t)$ exist for all $t > 0$ and in virtue of equation (2. 10) are equal for almost all $t > 0$. Here again by the use of (2. 11) we can show that both expressions are continuous and so the equality holds for all $t > 0$ and the boundary condition (2. 4) is satisfied for all $t > 0$.

By formal differentiation we have

$$(2.13)\quad \partial u_1/\partial r = -\int_0^t \{e^{-a^2(r-m)^2/4(t-t')} - e^{-a^2(r+m)^2/4(t-t')}\} r^{-2}(t-t')^{-1/2}\psi_1(t')dt'$$
$$-\int_0^t \{e^{-a^2(r-m)^2/4(t-t')} - e^{-a^2(r+m)^2/4(t-t')}\} a^2 2^{-1}(t-t')^{-3/2}\psi_1(t')dt'$$
$$+\int_0^t \{e^{-a^2(r-m)^2/4(t-t')} + e^{-a^2(r+m)^2/4(t-t')}\} a^2 m(2r)^{-1}(t-t')^{-3/2}\psi_1(t')dt'$$

with a similar expression for $\partial u_2/\partial r$; the integrals involved all converge, except possibly at $r = 0$, m or l. We have just shown that for $r \geq m/2$ the integrand in the first integral is less than a summable function and so the limit of the integral as r approaches m is equal to the integral of the limit of the integrand. The second and third integrals are of the type considered in the former paper where it is shown that †

$$\operatorname*{Limit}_{r=m\pm 0} (b/2\pi^{1/2})\int_0^t (x-m)(t-t')^{-3/2} e^{-b^2(x-m)^2/4(t-t')}\phi(t')dt' = \pm\phi(t)$$

almost everywhere for $\phi(t)$ summable. Applying this to equation (2. 13) gives

$$\operatorname*{Limit}_{r=m-0} \partial u_1/\partial r = -\int_0^t m^{-2}(t-t')^{-1/2}\{1 - e^{-a^2/(t-t')}\}\psi_1(t')dt'$$
$$+ a\pi^{1/2}m^{-1}\psi_1(t) + a^2\int_0^t (t-t')^{-3/2} e^{-a^2/(t-t')}\psi_1(t')dt'.$$

This holds for all $t > 0$. By forming the analogous expression for $\operatorname*{Limit}_{r=m+0} \partial u_2/\partial r$ we see that in virtue of equation (2. 9)

$$\operatorname*{Limit}_{r=m-0} K_1\partial u_1(r,t)/\partial r = \operatorname*{Limit}_{r=m+0} K_2\partial u_2(r,t)/\partial r, \quad \text{for} \quad t > 0,$$

* This remark was omitted from the former paper but applies there and is needed.

† Page 205.

nearly everywhere. Here again we show by use of (2. 11) that both sides are continuous and so the equation holds everywhere and the boundary condition (2. 5) is satisfied everywhere.

The solution given by the functions $u_1(r,t)$ and $u_2(r,t)$ defined by equations (2. 6) and (2. 7) in the inner and outer regions respectively thus satisfies the differential equations (2. 1), the initial conditions (2. 2) and the boundary conditions (2. 3) to (2. 5) nearly everywhere. We have already shown that it is bounded everywhere and must now show that the other conditions of the Uniqueness Theorem B are satisfied.

From the equations (2. 6) to (2. 7) we see that $u_1(r,t)$ and $u_2(r,t)$ are continuous for $t > 0$ except at $r = 0$, m and l, since the integral terms are the integrals of the product of a bounded, continuous function by a summable function.

The first derivatives can by a similar reasoning be shown to be continuous except at $r = 0$, m and l.

Finally we must show that

$$\operatorname*{Limit}_{r=r_0\pm 0} \int_{t_1}^{t_2} | \partial u_i(r,t)/\partial r \,|\, dt = \int_{t_1}^{t_2} \operatorname*{Limit}_{r=r_0\pm 0} | \partial u_i(r,t)/\partial r \,|\, dt, \quad \text{for} \quad t_1, t_2 > 0.$$

For $r \geq m/2$ the first integral in (2. 13) and in the similar expression for $\partial u_2/\partial r$ have been shown to be bounded and so satisfy the corresponding condition. The second and third integrals are of the form considered in the former paper where they were shown to satisfy the condition.* Thus $\partial u_1/\partial r$ and $\partial u_2/\partial r$ satisfy the condition.

If we multiply $\partial u_1/\partial r$ by r^2 each term is bounded and approaches zero with r and hence the same is true of the integral

$$\int_{t_1}^{t_2} r^2 \,| \partial u_1(r,t)/\partial r \,|\, dt.$$

Thus the solution given by $u_1(r,t)$ and $u_2(r,t)$ in the inner and outer regions, respectively, satisfies the conditions of the Uniqueness Theorem B.

It remains to be shown that summable functions $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$ can be found to satisfy the equations (2. 8) to (2. 10).

PART THREE. Solution of the Equations.

Precisely as before we derive from the equations (2. 8) to (2. 10) a set of integral equations for $\int \psi_1(t)dt$, $\int \psi_2(t)dt$ and $\int \psi_3(t)dt$. In the

* Page 207.

former problem these equations were Volterra integral equations of the second sort with bounded kernels. In this problem the equations again are Volterra equations of the second sort but the kernels are of the form

$$(t'' - t')^{-1/2} + N(t'', t')$$

where $N(t'', t')$ is bounded. Such a system is solvable by a process of successive approximations that necessarily converges and the solutions are bounded.*

As before we show that if the solutions of these equations are bounded, they are absolutely continuous and so possess derivatives, nearly everywhere, and are equal to the integrals of those derivatives.

By the same device previously employed we show that these derivatives—which are summable—satisfy the equations (2.8) to (2.10) and so give a solution of our problem.

* An elementary proof of this well-known fact is given by the author in a paper to appear in the *American Mathematical Monthly.*

HYPERELLIPTIC FUNCTIONS AND IRRATIONAL BINARY INVARIANTS. III.

By Arthur B. Coble.

Introduction. The second article [8] of this series was devoted to the linear irrational invariants (A) of the binary octavic, and to the irrational invariants (B) of the octavic of degree three. These invariants (B) may be regarded as the linear invariants of a special set of eight points in space, P_8^3, which are on a rational cubic norm-curve, N^3, say the invariants $d_{567} = (1234)(5678)$, the determinants being quaternary. The well-known linear identities among these 35 determinant products reduce to 14 the number which are linearly independent. Since even a generic set P_8^3 has only 9 absolute constants, the linear invariants must be subject to relations of higher degree. The definition of a set P_8^3 by means of its linear invariants is dependent upon the satisfaction of these relations of higher degree. Their determination is effected for the first time in section **12** below, there being 28 relations of the fifth degree in the d_{ijk}.

When these relations of the fifth degree are satisfied, and the d_{ijk} define a set P_8^3, it is clear that they must in general define two projectively distinct sets P_8^3; namely, a set P_8^3 and its associated set Q_8^3. For, complementary determinants formed from the coördinates of two associated sets are proportional, whence their linear invariants are proportional. However, the linear invariants define at most two projectively distinct sets P_8^3, Q_8^3. For, two sets Q_8^3, Q'^3_8, each associated with P_8^3, are projective to each other.

It may be observed that the sets P_8^3 are the first sets P^p_{2p+2} for which these relations of higher degree are present. When $p = 1$, the three linear invariants, $(12)(34)$, are linearly related. The ratio of any two, a double ratio of P_4^1, determines P_4^1 uniquely. In this case association implies projectivity. When $p = 2$, five of the ten linear invariants of P_6^2 are linearly independent. Their four ratios then define P_6^2 and its associated Q_6^2 [cf. [1], I, § 10].

In section **11** below we give some algebraic and geometric consequences of the determination of a pair of associated sets P_8^3, Q_8^3 by their linear invariants. These results will have analogues for all cases P^p_{2p+2}, Q^p_{2p+2} beyond $p = 1$.

If the associated sets P_8^3, Q_8^3 are also projective in the same order as they are associated, then the set P_8^3 is self-associated and is the set of base

points of a net of quadrics. Irrational conditions for such a P_8^3 are known [cf. [1], I, § 2 (20)], which lead to rational relations of the fourth degree. It is proved in **12** that, if these conditions for self-association are satisfied, then the quintic relations are also satisfied. Thus these conditions for self-association are sufficient to ensure the existence of a self-associated P_8^3. In **13** this self-associated P_8^3 is defined by a ternary set Q_7^2 with an attached quartic envelope, and the linear invariants of P_8^3 are expressed in terms of the Göpel invariants of the ternary quartic.

When the self-associated P_8^3 is subject to further conditions, also of the fourth degree, the P_8^3 becomes the hyperelliptic set P_8^3 on N^3. In **14** the linear invariants of this P_8^3 are developed in terms of the Göpel invariants of the underlying octavic. This section closes with the rational integral determination of the invariants (A) of the octavic, each multiplied by Δ, as polynomials of the fifth degree in the invariants (B).

11. Algebraic and geometric aspects of the linear invariants of a set of points, P_8^3. We set, as in **10** (1), (2), for the linear invariants,

$$d_{ijk} = \epsilon_{ijklmno}(ijk8)(lmno), \tag{1}$$

where $\epsilon_{ij\ldots o}$ is the sign of the permutation $ij\cdot\cdot\cdot o$ from the natural order $12\cdot\cdot\cdot 7$; and, for the 56 linear relations connecting them,

$$\begin{aligned} r_{12} &= d_{123} + d_{124} + d_{125} + d_{126} + d_{127} = 0, \\ r_{567} &= d_{567} + d_{234} + d_{134} + d_{124} + d_{123} = 0. \end{aligned} \tag{2}$$

Let us take the set P_8^3 with the last four points at the reference points. We write then the two arrays:

$$\begin{array}{cccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \cdot & \cdot & \cdot & \cdot \\ a_{41} & a_{42} & a_{43} & a_{44} \\ 1 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 \end{array} \qquad\qquad \begin{array}{cccc} A_{11} & A_{12} & A_{13} & A_{14} \\ \cdot & \cdot & \cdot & \cdot \\ A_{41} & A_{42} & A_{43} & A_{44} \\ -A & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & -A. \end{array} \tag{3}$$

In the second array we have the determinant $A = | a_{ij} |$, and the cofactors A_{ij} of the elements a_{ij}. Since each column of the one array has a zero product with each column of the other array, we have here two associated sets of eight points.

The 35 linear invariants of the first P_8^3 have a simple description. Each is, to within sign, a minor of A multiplied by its complementary minor, if we include $1, A$ as a pair of complementary minors. Thus the 18 pairs of

two-row minors, the 16 products $a_{ij}A_{ij}$, and the product $1 . A$, account for the 35 invariants. The 56 linear relations (2) comprise, in part, the Laplace expansions of A; in part, less familiar quadratic relations among the minors of A. We may state the theorem:

(4) *The necessary and sufficient conditions that* 35 *given constants may be the* 35 *products, each of a minor of a four-row determinant and its complementary minor, are that these constants shall satisfy the* 56 *linear relations* (2), *and the* 28 *quintic relations of* **12** (35). *When these are satisfied by the constants there are only two essentially distinct determinants which produce these constants and these may be taken as two adjoint determinants.*

It is understood here that a determinant is not essentially altered if a line is multiplied by a factor. This corresponds geometrically to the factor of proportionality in the coördinates of a point, and to a change in the unit point.

It is clear that the second array is that of the coördinates of the faces of the tetrahedra $p_1, \cdots, p_4; p_5, \cdots, p_8$. Hence

(5) *If a set P_8^3 of eight points in space is divided into two tetrahedra, the coördinates of the eight faces of the two tetrahedra are associated with the coördinates of the eight vertices.*

Since two sets each associated with a third are projective to each other, we can apply (5) to two different divisions to obtain

(6) *If P_8^3 is divided in two ways into two tetrahedra, the eight faces of the two tetrahedra in one division are projective to the eight faces of the two in the other division.*

From this projectivity between the two sets of eight faces, we get a projectivity between the two sets of eight vertices. This leads to a variety of theorems concerning projective relations among the points and the diagonal points of P_8^3 of which we give only the sample which arises from (6) for the two divisions 1234, 5678 and 1235, 4678. Consider the sets:

$$
\begin{array}{lllllllll}
(\alpha) & p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8; \\
(\beta) & \pi_{234} & \pi_{134} & \pi_{124} & \pi_{123} & \pi_{678} & \pi_{578} & \pi_{568} & \pi_{567}; \\
(\gamma) & \pi_{235} & \pi_{135} & \pi_{125} & \pi_{678} & \pi_{123} & \pi_{478} & \pi_{468} & \pi_{467}; \\
(\delta) & p_{15,678} & p_{25,678} & p_{35,678} & p_5 & p_4 & p_{46,123} & p_{47,123} & p_{48,123};
\end{array}
$$

where π_{ijk} is the plane $p_ip_jp_k$ and $p_{ij,klm}$ is the point where the line p_ip_j meets the plane $p_kp_lp_m$. According to (5) the points (α) are associated in order with both the planes (β) and the planes (γ). Hence the planes (β)

and (γ) are projective in order as in (6). Divide both (β) and (γ) into two tetrahedra made up of the first four and the last four planes. Then the eight vertices of the two tetrahedra (β), which are the points of (α) again, are projective to the eight vertices of the two tetrahedra (γ), which are the vertices given in (δ). Thus (α) and (δ) are projective in order. This projectivity may be described as follows:

(7) *If p_1, p_2, p_3 are projected from p_5 upon the plane $p_6p_7p_8$ to yield points q_1, q_2, q_3, and if p_6, p_7, p_8 are projected from p_4 upon the plane $p_1p_2p_3$ to yield points q_6, q_7, q_8, then $p_1, \cdots, p_8$ are projective in order to $q_1, q_2, q_3, p_5, p_4, q_6, q_7, q_8$.*

For different choices of the second division with respect to the first, and for different divisions of (β), (γ) into two tetrahedra, different projectivities are obtained.

12. The quintic relations satisfied by the linear invariants of P_8^3. For the purpose we have in mind it will be sufficient to examine further the quadratic irrational invariants of P_8^3. These are made up of four determinants $(ijkl)$, each point occurring in two determinants. An easy trial is sufficient to show that there exists but one type which is not a product of d_{ijk}'s, namely:

$$[12, 34;\ 56, 78] = (1257)(1268)(3458)(3467). \tag{1}$$

Of this type there are 35.6^2 exemplars corresponding to the 35. 6 pairs 12, 34, and to the six ways of matching two of the three combinations, 56, 78; 57, 68; 58, 67. The particular one given in (1) is invariant under the g_{32} generated by

$$(12),\ (34),\ (56)(78),\ (57)(68),\ (13)(24)(78). \tag{2}$$

We form the six invariants (1) for the particular choice 12, 34, and set

$$\begin{array}{ll} [12, 34;\ 67, 58] = a, & [12, 34;\ 76, 58] = \alpha, \\ [12, 34;\ 75, 68] = b, & [12, 34;\ 57, 68] = \beta, \\ [12, 34;\ 56, 78] = c, & [12, 34;\ 65, 78] = \gamma. \end{array} \tag{3}$$

If to c we apply in turn the identities,

$$(1257)(1268) + (1265)(1278) + (1276)(1258) = 0,$$
$$(3475)(3468) + (3456)(3478) + (3467)(3458) = 0,$$

we get $c + \beta - d_{125}d_{345} = 0$ and $c + \alpha - d_{126}d_{346} = 0$. The cyclic advance of 5, 6, 7 then yields:

$$(4)\qquad \begin{aligned} \lambda &\equiv d_{125}d_{345} = c + \beta = b + \gamma, \\ \mu &\equiv d_{126}d_{346} = a + \gamma = c + \alpha, \\ \nu &\equiv d_{127}d_{347} = b + \alpha = a + \beta. \end{aligned}$$

The six equations (4) for the determination of $a, b, c, \alpha, \beta, \gamma$ in terms of λ, μ, ν, i. e., in terms of the linear invariants, are dependent. The solution can be exhibited in terms of a new invariant of the second degree, $\delta_{12,34,5678}$, as follows:

$$(5)\qquad \begin{aligned} 2a &= \delta_{12,34,5678} - \lambda + \mu + \nu, & 2\alpha &= -\delta_{12,34,5678} - \lambda + \mu + \nu, \\ 2b &= \delta_{12,34,5678} + \lambda - \mu + \nu, & 2\beta &= -\delta_{12,34,5678} + \lambda - \mu + \nu, \\ 2c &= \delta_{12,34,5678} + \lambda + \mu - \nu, & 2\gamma &= -\delta_{12,34,5678} + \lambda + \mu - \nu, \end{aligned}$$

where

$$(6)\qquad \begin{aligned} \delta_{12,34,5678} &= a - \alpha = b - \beta = c - \gamma \\ &= (1257)(1268)(3458)(3467) - (1267)(1258)(3457)(3468). \end{aligned}$$

With this specific determination of the sign of $\delta_{12,34,5678}$ we see that

$$(7)\qquad \delta_{12,34,5678} = \delta_{21,34,5678} = \delta_{12,43,5678} = -\delta_{34,12,5678} = -\delta_{12,34,6578}.$$

From (5) and (3) we have

$$(8)\qquad \begin{aligned} a + \alpha &= -\lambda + \mu + \nu, & a\alpha &= \mu\nu, \\ b + \beta &= \lambda - \mu + \nu, & b\beta &= \nu\lambda, \\ c + \gamma &= \lambda + \mu - \nu, & c\gamma &= \lambda\mu. \end{aligned}$$

Hence

$$\begin{aligned} \delta^2_{12,34,5678} &= (a - \alpha)^2 = (a + \alpha)^2 - 4a\alpha \\ &= (-\lambda + \mu + \nu)^2 - 4\mu\nu. \end{aligned}$$

$$(9)\qquad \begin{aligned} \therefore\ \delta^2_{12,34,5678} &= \lambda^2 + \mu^2 + \nu^2 - 2\mu\nu - 2\nu\lambda - 2\lambda\mu \\ &= \Pi[\lambda^{1/2} + \mu^{1/2} + \nu^{1/2}]. \end{aligned}$$

From this there follows that

(10) *The invariants of the second degree of P_8^3 are all expressible in terms of the linear invariants and of the* 35. 6 *invariants $\delta_{ij,kl,mnop}$ of the second degree. The squares of these invariants δ are expressible as quartic polynomials in the linear invariants. The vanishing of the invariants δ is the necessary and sufficient condition that P_8^3 be self-associated.*

For, the irrational condition, $\lambda^{1/2} + \mu^{1/2} + \nu^{1/2} = 0$, expresses that the two sets of four planes on the lines p_1p_2 and p_3p_4 respectively to the four points p_5, p_6, p_7, p_8 are projective. These conditions are sufficient to ensure self-association [cf. [1], I, p. 165 (20)]. The three alternative forms of $-\delta_{12,34,5678}$ in (6) are the three determinants of the matrix

$$(11) \qquad \left\| \begin{matrix} (1256)(1278) & (1257)(1286) & (1258)(1267) \\ (3456)(3478) & (3457)(3486) & (3458)(3467) \end{matrix} \right\| .$$

Since the sum of the elements in each row is zero, the three determinants are equal. The vanishing of a particular determinant evidently expresses the projective situation just mentioned. We shall find in the next two sections that these irrational conditions are merely a version of the three-term linear relations connecting the Göpel invariants.

That the invariants δ vanish for a self-associated set may be seen in (11). For, the proportionality of complementary determinants of such a set, expressed by

$$(12) \qquad (ijkl) = \epsilon_{ijklmnop} \cdot \lambda \cdot (mnop) \qquad [\lambda^2 = 1],$$

where ϵ is the sign of the attached permutation [cf. [1], I, p. 158 (7)], shows that the determinants in (11) are unaltered if the rows are interchanged.

The invariants δ are connected by a system of three-term linear relations. These are derived from a four-term determinant identity as follows.

$$\begin{aligned} &[12, 34;\ 56, 78] + [12, 45;\ 36, 78] + [12, 53;\ 46, 78] \\ &= (1268)(3458)\{(1257)(3467) + (1237)(4567) + (1247)(5367)\} \\ &= (1268)(3458)(1267)(5347) = -d_{126}d_{345}. \end{aligned}$$

In this identity interchange 7, 8 and subtract, making use of the definition (6) of the invariants δ. Then

$$(13) \qquad \delta_{12,34,5678} + \delta_{12,45,3678} + \delta_{12,53,4678} = 0.$$

Recollecting that, according to (9), the squares of the δ's are polynomials of the fourth degree in the linear invariants we write (13) in the irrational form

$$(14) \qquad \{\delta^2_{12,34,5678}\}^{1/2} + \{\delta^2_{12,45,3678}\}^{1/2} + \{\delta^2_{12,53,4678}\}^{1/2} = 0.$$

On rationalizing this relation we secure an octavic relation among the linear invariants from which the product $d_{345}d_{126}d_{127}$ may be divided out leaving a quintic relation among the linear invariants. For, if we set, after the pattern of (4),

$$(15.1) \qquad \begin{matrix} \lambda_1 = d_{125}d_{345}, & \lambda_2 = d_{123}d_{345}, & \lambda_3 = d_{124}d_{345}, \\ \mu_1 = d_{126}d_{346}, & \mu_2 = d_{126}d_{456}, & \mu_3 = d_{126}d_{536}, \\ \nu_1 = d_{127}d_{347}, & \nu_2 = d_{127}d_{457}, & \nu_3 = d_{127}d_{537}, \end{matrix}$$

the rationalized relation is

$$(15.2) \qquad \sum_{i=1}^{i=3} \{\lambda_i^2 + \mu_i^2 + \nu_i^2 - 2\mu_i\nu_i - 2\nu_i\lambda_i - 2\lambda_i\mu_i\}^{1/2} = 0.$$

The terms which do not contain d_{345} explicitly have the form

$$\{(\mu_1-\nu_1)^2\}^{1/2}+\{(\mu_2-\nu_2)^2\}^{1/2}+\{(\mu_3-\nu_3)^2\}^{1/2}.$$

But

$$\begin{aligned}(\mu_1-\nu_1)&+(\mu_2-\nu_2)+(\mu_3-\nu_3)\\ &=d_{126}(d_{346}+d_{456}+d_{356})-d_{127}(d_{347}+d_{457}+d_{357})\\ &=d_{126}(-d_{345}-d_{127})+d_{127}(d_{345}+d_{126})=d_{345}(d_{127}-d_{126}).\end{aligned}$$

Thus d_{345}, and similarly d_{126} and d_{127}, can be isolated from the octavic relation leaving a quintic relation connecting the d_{ijk}.

In the examination of this quintic relation, and of others to be derived later, we shall at first be concerned only with those linear invariants in which the pair 12 occurs in a determinant of the product. For these we shall introduce a special notation as follows:

$$(16)\qquad x_{i8}=d_{12i},\qquad x_{ij}=d_{klm}\qquad (i,j,\cdots,m=3,\cdots,7).$$

Thus the 15 linear invariants retained x_{ij} $(i,j=3,\cdots,8)$, are connected by six linear relations of the form

$$(17)\qquad x_{ij}+x_{ik}+x_{il}+x_{im}+x_{in}=0\qquad (i,\cdots,n=3,\cdots,8).$$

Five of these may be used to eliminate $x_{34},\cdots,x_{38}$, leaving ten invariants x_{ij} which are connected by the single relation [cf. [1], I, p. 188],

$$(18)\qquad \Sigma\, x_{jk}=0\qquad (j,k=4,\cdots,8).$$

In this notation the products λ, μ, ν associated above with the quadratic invariant $\delta_{12,34,5678}$ are

$$(19)\qquad x_{58}x_{67},\quad x_{68}x_{57},\quad x_{78}x_{56}.$$

The relation (15) now reads

$$(20)\qquad \sum_{i=3}^{i=5}\{x^2_{i6}x^2_{78}+x^2_{i7}x^2_{68}+x^2_{i8}x^2_{67}-2x_{i7}x_{i8}x_{68}x_{67}\\ \qquad\qquad -2x_{i8}x_{i6}x_{67}x_{68}-2x_{i6}x_{i7}x_{78}x_{68}\}^{1/2}=0.$$

From the rationalized form the factor $x_{67}x_{68}x_{78}$ must separate as has been pointed out. If the terms containing a subscript 3 are replaced in terms of the others as in (17), an elementary calculation yields, after deleting the factor 16 $x_{67}x_{68}x_{78}$, the following form of the quintic relation:

$$(21)\quad Q_{12}\equiv\Sigma\, x^2_{45}x_{67}x_{68}x_{78}-\Sigma(x_{45}x_{56}x_{67}x_{78}x_{84}+x_{46}x_{68}x_{85}x_{57}x_{74})=0,$$

where the summations are symmetric in the subscripts $4,\cdots,8$, the first being over ten terms, and the second over the six terms determined by the

six cyclic g_5's on the five subscripts. The relation sought may however take alternative forms due to the linear relation (18) among the invariants. The particular form (21) is obtained by using the following forms of the δ^2's, and of the relation:

$$(22)\quad \begin{aligned} \delta^2_{12,34,5678} &= (x_{67}x_{58} - x_{57}x_{68})^2 - 2x_{56}x_{78}(x_{67}x_{58} + x_{57}x_{68}) + x^2_{56}x^2_{78}, \\ \delta^2_{12,45,3678} &= (x_{68}x_{47} - x_{48}x_{67} - x_{67}x_{58} + x_{57}x_{68})^2 + 2\{2x_{45}x_{67}x_{68} \\ &\quad - (x_{56} + x_{64})(x_{68}x_{47} + x_{48}x_{67} + x_{67}x_{58} + x_{57}x_{68}\}x_{78} + (x_{56} + x_{64})^2x^2_{78}. \end{aligned}$$

With the terms in x_{3i} now deleted from the relation,

$$(\delta^2_{12,34,5678} - \delta^2_{12,35,4678})^2 - 2\delta^2_{12,45,3678}(\delta^2_{12,34,5678} + \delta^2_{12,35,4678}) + \delta^4_{12,45,3678} = 0,$$

and with terms arranged in powers of x_{78}, the terms in x^0_{78} and x^4_{78} vanish, and the form (21) appears.

The polynomial Q_{12} obviously admits a $g_{5!}$ due to its symmetry in the indices $4, \cdots, 8$. We prove now that it actually is symmetric in the six indices $3, \cdots, 8$ and thus admits a $g_{6!}$. If the subscripts 3, 4 be interchanged to produce the collineation,

$$(23)\quad \begin{aligned} x'_{45} &= x_{35} = -x_{45} \phantom{- x_{56}} - x_{56} - x_{57} - x_{58}, \\ x'_{46} &= x_{36} = -x_{46} - x_{56} \phantom{- x_{67}} - x_{67} - x_{68}, \\ x'_{47} &= x_{37} = -x_{47} - x_{57} - x_{67} \phantom{- x_{68}} - x_{78}, \\ x'_{48} &= x_{38} = -x_{48} - x_{58} - x_{68} - x_{78}, \\ x'_{ij} &= x_{ij} \qquad (i, j = 5, \cdots, 8), \end{aligned}$$

we have to show only that this linear transformation leaves unaltered both the linear relation (18) and the form Q_{12}. Evidently the transformation converts (18) into its negative and the relation is unaltered.

The transformation may of course be applied directly to the form Q_{12}, but we prefer to prove the invariance of Q_{12} by exhibiting some properties of the quintic manifold $Q_{12} = 0$ in the S_8 defined by the ten linearly related variables. It is clear from the equation (21) that the point, $x_{45} = 1$, $x_{46} = -1$ (the other coördinates being zero), is a triple point of Q_{12}, and that the tangent cone at the triple point is

$$(24)\quad [x_{45} : x_{46} = 1 : -1] \quad 2x_{78}[(x_{57} + x_{67})(x_{58} + x_{68}) - x_{56}x_{78}] = 0.$$

The manifold Q_{12} has 30 triple points of this character of the form $x_{ij} : x_{ik} = 1 : -1$ $(i, j, k = 4, \cdots, 8)$. These are a conjugate set under the collineation $g_{5!}$ induced by permutation of the indices $4, \cdots, 8$. The collineation (23) transforms $x_{45} : x_{46} = 1 : -1$ into itself; it transforms $x_{45} : x_{56} = 1 : -1$ into $x_{46} : x_{56} = 1 : -1$; and it transforms $x_{56} : x_{57} = 1 : -1$ into $x_{46} : x_{47} : x_{56} : x_{57} = -1 : 1 : 1 : -1$. The number of

points of this character under the collineation group $g_{5!}$ is 15. We denote the last point by the symbol $P_{45,67}$ and its conjugate $x_{56} : x_{57} = 1 : -1$ by $P_{35,67}$. Since (23) also transforms $P_{56,78}$ into itself, the 45 points $P_{ij,kl}$ $(i, \cdots, l = 3, \cdots, 8)$ are a conjugate set under the collineation $g_{6!}$ induced by permutation of the indices $3, \cdots, 8$.

The point $P_{45,67}$ is also a triple point of Q_{12}. For, the only terms which persist in the quadric polar have coefficients which cancel. The tangent cone at the triple point is

$$2(x_{68} + x_{78})\{x_{48} + x_{58})(x_{57} + x_{56} + x_{47} + x_{46}) - x_{45}(x_{68} + x_{78})\} \\ - 2x_{67}(x_{58} + x_{48})^2.$$

If the expression $(x_{45} + x_{67})(x_{48} + x_{58})$ be added and subtracted from this, the factor $(x_{48} + x_{58} + x_{68} + x_{78})$ appears, and the cone takes the form

$$2x'_{48}\{(x'_{46} + x'_{47})(x'_{68} + x'_{78}) - x'_{67}x'_{48}\} \qquad \text{[cf. (23)]}.$$

On comparing this with (24) we see that the collineation (23) has interchanged the triple points $P_{35,67}$, $P_{45,67}$, and also has interchanged their tangent cones. It is not difficult to verify that this is true of each triple point and tangent cone. Typical cases are $P_{34,56}$, $P_{35,46}$, $P_{35,67}$ (just discussed), and $P_{56,78}$. Furthermore it is clear that the manifold Q_{12} is defined by its set of 30 triple points $P_{3i,jk}$ and their associated tangent cones. Since this set passes under the collineation (23) into a similar set of 30 triple points and associated tangent cones of the *same* manifold Q_{12}, there follows that:

(25) *In the S_8 defined by the variables x_{ij} $(i, j = 3, \cdots, 8)$ the quintic manifold $Q_{12} = 0$ is invariant under the collineation group $g_{6!}$ induced by permutation of the indices. It is characterized by the existence of a conjugate set of 45 triple points $P_{ij,kl}$.*

It should be noted that the 28 relations $Q_{ij} = 0$ $(i, j = 1, 2, \cdots, 8)$ are necessary conditions on the linear invariants d_{ijk} that they may define a set of eight points in space. For, they have been obtained on the hypothesis that the quadratic invariants δ exist in connection with such a set. We seek now to prove that these 28 relations are sufficient conditions that P_8^3 may exist for the given d_{ijk}.

We first obtain expressions for the double ratios in pencils of planes on lines of P_8^3. Denoting by $D(56, 78)$ the usual binary double ratio $(57)(68)/(58)(67)$, and by $D(12; 56, 78)$ the corresponding double ratio of the four planes on the line p_1p_2 to the four points $p_5, \cdots, p_8$, then

$$(26)\quad \begin{aligned} D(12;\,56,78) &= (1257)(1268)/(1258)(1267) \\ &= (1257)(1268)(3456)(3478)/(1258)(1267)(3456)(3478) \\ &= -[12,34;\,76,58]/[12,34;\,75,68] \\ &= \frac{\delta_{12,34,5678}+x_{58}x_{67}-x_{68}x_{57}-x_{78}x_{56}}{\delta_{12,34,5678}+x_{58}x_{67}-x_{68}x_{57}+x_{78}x_{56}} \end{aligned}$$

[cf. (1), (5), (16)]. Now the double ratios formed from a set of more than four points on a line are subject to a system of cubic relations [cf. **9** (2)] which ensure their consistency in the determination of the set. These have the form

$$D(56,78)\cdot D(64,78)\cdot D(45,78)=1.$$

On applying this to (26) we have the relation

$$(27)\qquad \Pi\,\frac{\delta_{12,34,5678}+x_{58}x_{67}-x_{68}x_{57}-x_{78}x_{56}}{\delta_{12,34,5678}+x_{58}x_{67}-x_{68}x_{57}+x_{78}x_{56}}=1,$$

the product Π being formed for the cyclic advance of 4, 5, 6. The system of relations of type (27) ensures the existence of sextics of planes, each associated with a line joining two of the points of the hypothetical P_8^3.

If P_8^3 exists, and the first five of its points are taken at the reference points and unit point, then the sets of six planes on the lines p_1p_2, p_1p_3, p_2p_3, whose existence is assured as a consequence of (27), will determine the position of p_6, p_7, p_8. We have in fact for p_8 the coördinates,

$$D(12;\,34,58)=p_{8,3}/p_{8,4},\qquad D(23;\,14,58)=p_{8,1}/p_{8,4},$$
$$D(31;\,24,58)=p_{8,2}/p_{8,4}.$$

But this position of p_8 as determined from these three pencils must be consistent with its position as determined from an adjacent pencil, e. g., that on p_1p_4. But $D(14;\,23,58)=p_{8,2}/p_{8,3}$. Hence another type of cubic relation among the double ratios appears, namely:

$$(28)\qquad D(12;\,43,58)\cdot D(13;\,24,58)\cdot D(14;\,32,58)=1.$$

We express this first in terms of the quadratic invariants as follows:

$$\Pi_{2,3,4}[67,12;\,35,48]/[67,12;\,45,38]=-1;$$

and then pass to a form comparable with (27) by applying the permutation (1357246). The result, expressed in terms of the δ's, is

$$(29)\qquad \Pi_{4,5,6}\,\frac{\delta_{12,34,5678}-x_{58}x_{67}+x_{68}x_{57}-x_{78}x_{56}}{\delta_{12,34,5678}-x_{58}x_{67}+x_{68}x_{57}+x_{78}x_{56}}=1.$$

From the method of derivation of the relations (27) and (29) there follows that

(30) *The existence of the relations* (27) *and* (29) *on the linear invariants* d_{ijk} *are the necessary and sufficient conditions that the* d_{ijk} *belong to a set* P_8^3.

We wish now to prove that all of these relations are consequences of the 28 relations Q_{ij} of type (21). In (27) and (29) each factor of the numerator is identical with the corresponding factor of the denominator except for the change of sign in the last term, and all three of the last terms have the common factor x_{78} which therefore is a factor in the relation. After separating this factor the relation (27) takes the form

$$(31) \qquad x^2_{78}x_{56}x_{64}x_{45} + \Sigma_{4,5,6}x_{56}(\delta_{12,35,6478} + x_{68}x_{47} - x_{48}x_{67})(\delta_{12,36,4578} + x_{48}x_{57} - x_{58}x_{47}) = 0.$$

On the other hand, the relation similarly deduced from (29) differs from (31) only by a change of sign equivalent to a change of sign of the δ's. Hence, on adding and subtracting the simplified relations, we obtain two equivalent types, namely:

$$(32.1) \qquad x^2_{78}x_{56}x_{64}x_{45} + \Sigma_{4,5,6}x_{56}\{\delta_{12,35,6478}\delta_{12,36,4578} + (x_{68}x_{47} - x_{48}x_{67})(x_{48}x_{57} - x_{58}x_{47})\} = 0,$$

$$(32.2) \qquad \Sigma_{4,5,6}\delta_{12,34,5678}\{x_{64}(x_{48}x_{57} - x_{58}x_{47}) + x_{45}(x_{68}x_{47} - x_{48}x_{67})\} = 0.$$

The first of these two, having terms of even degree only in the δ's, is a polynomial in the d_{ijk}. For, if a term of (13) be transposed, the equal squares yield a formula which, for the above case, reads

$$(33) \qquad 2\delta_{12,35,6478}\delta_{12,36,4578} = \delta^2_{12,35,6478} + \delta^2_{12,36,4578} - \delta^2_{12,56,3478}.$$

On applying this formula to the product of (13) by one of its terms, we obtain another relation of the type:

$$(34) \qquad 2\delta_{12,34,5678}\delta_{12,56,3478} = \delta^2_{12,36,4578} + \delta^2_{12,45,3678} - \delta^2_{12,35,4678} - \delta^2_{12,46,3578}.$$

If the δ products in (32.1) are replaced by squares from (33), and these in turn are expressed as in (22), the relation (32.1) takes the form $-4Q_{12} = 0$ [cf. (21)].

Furthermore, if the relation (32.2) be multiplied by $\delta_{12,34,5678}$, and the expressions be modified as before, it takes the form

$$-4(x_{58}x_{67} - x_{68}x_{57}) \cdot Q_{12} = 0.$$

Hence

(35) *The relations* (27), (29), *and the equivalent relations* (32.1), (32.2) *all subsist by virtue of the* 28 *relations* $Q_{ij} = 0$ $(i, j = 1, \cdots, 8)$. *These* 28 *relations of the fifth degree in the linear invariants* d_{ijk} *are the necessary*

and sufficient conditions that these quantities d_{ijk} may be the linear invariants of a set P_8^3 of 8 points in space.

Let us set, for the moment,

$$\delta^2_{ij} = \delta^2_{12,ij,klmn} \qquad (i, j = 3, \cdots, 8). \tag{36}$$

Then, according to (22), the $\delta^2_{34} = 0$ has a 4-fold point at $P_{34,56}$ and a double point at $P_{35,46}$, $P_{35,67}$, and $P_{56,78}$. Hence $x_{34}\delta^2_{34} = 0$ has a 5-fold point at $P_{34,56}$, 3-fold points at $P_{35,67}$ and $P_{56,78}$, and a double point only at $P_{35,46}$. Thus $P_{34,56}$ is at least a triple point on every $x_{ij}\delta^2_{ij} = 0$ except those in $x_{35}\delta^2_{35} + x_{36}\delta^2_{36} + x_{45}\delta^2_{45} + x_{46}\delta^2_{46}$. If, however, we examine this sum, it appears at once that the terms which vanish only doubly at $P_{34,56}$ cancel each other. Hence the sum $\Sigma\, x_{ij}\delta^2_{ij}$ has triple points at all of the points $P_{ij,kl}$, and we prove that

$$10Q_{12} = \Sigma_{i,j} x_{ij}\delta^2_{12,ij,klmn} \qquad (i, \cdots = 3, \cdots, 8). \tag{37}$$

The proof consists merely in showing that the tangent cone of the sum at $P_{34,56}$, obtained by operation with $\partial^2/\partial x^2_{45} - 2\partial^2/\partial x_{45}\partial x_{46} + \partial^2/\partial x^2_{46}$, is 10 times the tangent cone of Q_{12} at $P_{34,56}$, this being given in (24). Since Q_{12} has been shown to be invariant under the collineation $g_{6!}$ of the indices, and the sum is likewise invariant under $g_{6!}$, then both members of (37) have the same triple points and respective tangent cones at these points, and therefore both are the same. We omit the elementary identification of these tangent cones and merely give the result (37).

The formula (37), interpreted in the light of theorems (35) and (10), yields this result:

(38) *If the linear invariants d_{ijk} satisfy the relations of the fourth degree of the type $\delta^2_{ij,kl,mnop} = 0$ $(i, \cdots = 1, \cdots, 8)$, then they are the linear invariants of a self-associated set P_8^3, the base points of a net of quadrics.*

The determination of a set P_8^3 for given d_{ijk} subject to the quintic relations of (35) would proceed as follows. Since $\delta^2_{12,34,5678}$ is a polynomial of degree four in d_{ijk}, we take one of the two values of $\delta_{12,34,5678}$, and thereby select one of the two associated P_8^3's which have the same values of d_{ijk}. The sign of every $\delta_{ij,kl,mnop}$ is then uniquely determined. For, by repeated applications of (33) and (34) whose right members are polynomials of degree four in the d_{ijk}, the sign of every $\delta_{12,ij,klmn}$ is determined by the choice of sign of $\delta_{12,34,5678}$. But, according to (7), $\delta_{12,ij,klmn} = -\delta_{ij,12,klmn}$, and thereby the sign of $\delta_{ij,kl,mnop}$ $(i, j \neq 1, 2)$ is determined. Again by the use of (33) and (7), the sign of $\delta_{ij,kl,mnop}$, when either k or l is 1 or 2, or when

i or j is 1 or 2, is likewise determined. Then, by using formulae (5), the values of the invariants $[ij, kl; mn, op]$ are obtained. From these values, the double ratios of the set P_8^3 are obtained as in (26). When the δ's all vanish, and the set is self-associated, these double ratios are given at once by (26) as the ratios of polynomials quadratic in the d_{ijk}'s.

13. The self-associated set P_8^3 which is not on a cubic space curve. The generic set of eight base points of a net of quadrics is distinguished by the fact that any eighth point of it is uniquely determined when seven are given. If these seven are $P_7^3 = p_1, \cdots, p_7$, the space set P_7^3 is associated with a planar set $Q_7^2 = q_1, \cdots, q_7$, which is, if we please, the projection of the set P_7^3 from the eighth base point p_8. The set Q_7^2 being given, the associated P_7^3 is projectively determined, and thereby the self-associated P_8^3 is projectively determined. There would be therefore an obvious advantage in expressing the irrational invariants of the self-associated P_8^3 in terms of the invariants of Q_7^2, since the points of the latter set are generic, while the points of P_8^3 are conditioned.

The 63 discriminant conditions attached to P_8^3 comprise 28 of type ϵ_{ij} which indicate coincidence of two points in some direction; and 35 of type $\epsilon_{ijkl} = \epsilon_{mnop}$ which are the coplanar conditions of the set, the equality being a consequence of self-association. The 63 discriminant conditions attached to Q_7^2 comprise 21 of type δ_{ij} which indicate coincidence, 35 of type δ_{ijk} which indicate collinearity, and 7 of type δ_{i8} which indicate that the points of Q_7^2 other than p_i are on a conic. These are the discriminant factors of the quartic envelope with nodes at Q_7^2, and also the discriminant factors of the birationally equivalent sextic locus of nodes of quadrics on P_8^3. They are related as follows:

$$(1) \qquad \delta_{ij} = \epsilon_{ij}, \quad \delta_{i8} = \epsilon_{i8}, \quad \delta_{ijk} = \epsilon_{ijk8} = \epsilon_{mnop}.$$

We fix the signs of the 7 discriminant conditions, $\Delta_i = \delta_{i8} = 0$, by writing the identity connecting the squares of the seven points of Q_7^2 in the form

$$(2) \qquad \Delta_1(q_1\eta)^2 + \Delta_2(q_2\eta)^2 + \cdots + \Delta_7(q_7\eta)^2 = 0.$$

We make this more precise by the definition:

$$(3) \qquad \Delta_7 = (134)(156)(253)(246) - (234)(256)(153)(146).$$

Then a Δ_i obtained from this by a permutation is $+\Delta_i$ or $-\Delta_i$ according as the permutation is even or odd. If η in (2) is the line q_6q_7, we have

$$(4) \qquad \Delta_1(167)^2 + \Delta_2(267)^2 + \cdots + \Delta_5(567)^2 = 0;$$

if (2) is polarized as to η, ζ and η is q_6q_7, ζ is q_4q_5, then

$$(5) \qquad \Delta_1(145)(167)+\Delta_2(245)(267)+\Delta_3(345)(367)=0.$$

The linear invariant, $d_{567}=(1234)(5678)$, of P_8^3 vanishes with the 14 discriminant conditions, $\epsilon_{12},\cdots,\epsilon_{34},\epsilon_{56},\cdots,\epsilon_{78},\epsilon^2_{5678}$. We examine the planar product $\Delta_5\Delta_6\Delta_7(567)^2$. This product is of the sixth degree in each point of Q_7^2, and it vanishes twice for each coincidence δ_{ij}. It vanishes triply for the coincidences $\delta_{12},\cdots,\delta_{34},\delta_{56},\delta_{57},\delta_{67}$, and it vanishes simply for $\delta_{58},\delta_{68},\delta_{78}$, and doubly for δ_{567}. We therefore are led to set

$$(6) \qquad d_{ijk}=\Delta_i\,\Delta_j\,\Delta_k\,(ijk)^2 \qquad (i,j,k=1,\cdots,7).$$

This will be justified if we show that the right members satisfy the same linear relations, and the same quartic relations (those which ensure self-association) as the left members. But, if (4) is multiplied by $\Delta_6\Delta_7$ it yields, in accordance with (6), the linear relation, $d_{167}+d_{267}+\cdots+d_{567}=0$, which characterizes the d_{ijk}. If also (5) is multiplied by $(\Delta_4\Delta_5\Delta_6\Delta_7)^{1/2}$, it yields the irrational condition,

$$(7.1) \qquad (d_{156}d_{167})^{1/2}+(d_{245}d_{267})^{1/2}+(d_{345}d_{367})^{1/2}=0,$$

satisfied by the self-associated P_8^3. Hence

(7.2) *For given generic Q_7^2 the equations* (6) *define the linear invariants, d_{ijk}, of the self-associated P_8^3 whose points $p_1,\cdots,p_7$ are associated with Q_7^2, i. e., which, projected from p_8, are projective to Q_7^2.*

The Göpel invariants of the generic Q_7^2 are of the third degree in the coördinates of each point q_i and vanish at least once for each coincidence δ_{ij}. Their explicit values are obtained from

$$(8.1) \qquad \begin{aligned} [cf+] &= 8(547)(217)(367)\cdot(531)(461)(342)(562),\\ [cf-] &= 8(547)(217)(367)\cdot(523)(462)(341)(561),\\ [cf] &= 8(547)(217)(367)\cdot\Delta_7,\\ [be,cf] &= -[be-]-[cf+], \end{aligned}$$

by using the parallel permutations of **8** (2). These 135 Göpel invariants of Q_7^2 satisfy the 315 three-term relations of **8** $(\alpha),\cdots,(\zeta)$ [cf. also [1], II, pp. 380-384; [3], pp. 192-197]. The right members of (6) are of degree 6 in the coördinates of each point of Q_7^2, and they vanish at least twice for each coincidence. It may be expected therefore that they should be of degree two in the Göpel invariants. We get by permutation from (8.1)

$$
\begin{aligned}
&[ca, fd][ca, ef] - [bc, de][ab, de] \\
&\qquad = 64\{\Delta_3(357)(321)(346)\cdot\Delta_2(267)(245)(231) \\
&\qquad\qquad - \Delta_3(367)(345)(312)\cdot\Delta_2(257)(213)(246)\} \\
&\qquad = 64\Delta_1\,\Delta_2\,\Delta_3(123)^2 = 64 d_{123}.
\end{aligned}
$$

Similarly we find that

$$
\begin{aligned}
&[bf][bf, ce] - [ad][ce, ad] \\
&\qquad = 64\{\Delta_7(327)(147)(657)\cdot\Delta_5(567)(534)(512) \\
&\qquad\qquad - \Delta_7(657)(127)(437)\cdot\Delta_5(567)(541)(532)\} \\
&\qquad = 64\Delta_5\,\Delta_6\,\Delta_7(567)^2 = 64 d_{567}.
\end{aligned}
$$

Making use of the three-term relations among the Göpel invariants, and of the fact that all of the cofactors of a three-row determinant are equal if the sum of the elements in each line is zero, we have the theorem:

(8.2) *The linear invariants, $64d_{123}$, $64d_{567}$, of the self-associated $P_8{}^3$ are equal to the cofactors in the respective determinants,*

$$
\begin{vmatrix} [ca, fd] & [ab, de] & [bc, ef] \\ [bc, de] & [ca, ef] & [ab, fd] \\ [ab, fe] & [bc, fd] & [ca, de] \end{vmatrix}, \qquad
\begin{vmatrix} [bf, ce] & [ce, bf] & [ad] \\ [ad, bf] & [bf, ad] & [ce] \\ [ce, ad] & [ad, ce] & [bf] \end{vmatrix};
$$

and thus are of the second degree in the Göpel invariants of the planar $Q_7{}^2$.

The two invariants given above represent conjugate sets of 20 and 15 respectively under the parallel permutations of $1, \cdots, 6$ and of $a, \cdots, f$. The ternary quartic has a conjugate set of 630 such invariants. For, it defines 36 self-associated $P_8{}^3$'s, and each has 35 linear invariants. Each invariant however belongs to two $P_8{}^3$'s. For example, d_{567} belongs both to the given $P_8{}^3$ and to that which arises from it by the cubic Cremona transformation with four double F-points at $p_5, \cdots, p_8$ (or at $p_1, \cdots, p_4$) [cf. [1], II (45)].

Since there are 15 linearly independent Göpel invariants, there are 120 independent quadratic combinations. There being 135 Göpel invariants, their 135 squares must be connected by 15 linear relations. We proceed to find these relations, and to find expressions for, not merely the d_{123}, d_{567} in (8.2), but also other significant products of the d_{ijk}, in terms of these Göpel squares. For brevity we set

(9.1) $$x_{ab} \equiv [ab\,-], \quad y_{ab} \equiv [ab\,+],$$

(9.2) $$\therefore \quad [ab] = -x_{ab} - y_{ab}, \quad [ab, cd] = -x_{ab} - y_{cd}.$$

The remaining three-term linear relations among the Göpel invariants are now all comprised under the following two sets of 15 and 20 respectively:

(10. 1) $$x_{ab}+x_{cd}+x_{ef}+y_{ab}+y_{cd}+y_{ef}=0,$$

(10. 2) $$x_{ab}+x_{ac}+x_{bc}+y_{de}+y_{df}+y_{ef}=0.$$

If we set

(11. 1) $$\sigma_x=\Sigma_{15}x_{ab},\qquad \sigma_y=\Sigma_{15}y_{ab},$$

(11. 2) $$\sigma_{x^2}=\Sigma_{15}x^2{}_{ab},\qquad \sigma_{y^2}=\Sigma_{15}y^2{}_{ab},$$

then we find from (10) that

(12) $$\sigma_x+\sigma_y=0,\qquad \sigma_{x^2}-\sigma_{y^2}=0.$$

If (10. 2) be formed for *abc, abd, abe, abf* in y, and if the results be added, we obtain (13. 1), and similarly (13. 2) where

(13. 1) $$3y_{ab}+\sigma_y-\Sigma_6 y_{cd}=-2\Sigma_6 x_{cd},$$

(13. 2) $$3x_{ab}+\sigma_x-\Sigma_6 x_{cd}=-2\Sigma_6 y_{cd}.$$

By eliminating $\Sigma_6 y_{cd}$ we find that

(14. 1) $$6[ab-]=6x_{ab}=6x_{ab},$$

(14. 2) $$6[ab+]=6y_{ab}=\sigma_x-3x_{ab}-3\Sigma_6 x_{cd},$$

(14. 3) $$6[ab]=-6(x_{ab}+y_{ab})=-\sigma_x-3x_{ab}+3\Sigma_6 x_{cd},$$

(14. 4) $$6[cd,ab]=-6(x_{cd}+y_{ab})=-\sigma_x+3x_{ab}+3\Sigma_6 x_{cd}-6x_{cd}.$$

Thus the 135 Göpel invariants are expressed linearly in terms of the 15 x_{ab}'s which themselves are independent. The similar expressions in terms of the 15 y_{ab}'s are obvious. The Göpel invariants defined for the hyperelliptic case in part II [cf. [8]] satisfy precisely similar linear relations except that an additional linear relation $\sigma_x=\sigma_y=0$ replaces $\sigma_x+\sigma_y=0$ in (12).

The squares of the formulae (9. 2) yield

(15. 1) $$[ab]^2=x^2{}_{ab}+y^2{}_{ab}+2x_{ab}y_{ab},$$

(15. 2) $$[ab,cd]^2=x^2{}_{ab}+y^2{}_{cd}+2x_{ab}y_{cd},$$

whence

(16) *The* 135 *products* $x^2{}_{ab}$, $y^2{}_{ab}$, $x_{ab}y_{cd}$, $x_{ab}y_{ab}$ *can all be expressed linearly in terms of the Göpel squares and conversely.*

Other products can be expressed linearly in terms of these 135 products

with some simplicity. Thus if (10.2) be multiplied in turn by x_{bc}, x_{ca}, x_{ab} we find that

$$(17.1)\qquad 2x_{ac}x_{bc}=x^2_{ab}-x^2_{ac}-x^2_{bc}+(x_{ab}-x_{ac}-x_{bc})(y_{de}+y_{df}+y_{ef}),$$

$$(17.2)\qquad \therefore\ 2y_{ac}y_{bc}=y^2_{ab}-y^2_{ac}-y^2_{bc}+(y_{ab}-y_{ac}-y_{bc})(x_{de}+x_{df}+x_{ef}).$$

If (10.1) be multiplied in turn by x_{ab}, x_{cd}, x_{ef} we find that

$$(18.1)\qquad 2x_{ab}x_{cd}=x^2_{ef}-x^2_{ab}-x^2_{cd}+(x_{ef}-x_{ab}-x_{cd})(y_{ef}+y_{ab}+y_{cd}),$$

$$(18.2)\qquad \therefore\ 2y_{ab}y_{cd}=y^2_{ef}-y^2_{ab}-y^2_{cd}+(y_{ef}-y_{ab}-y_{cd})(x_{ef}+x_{ab}+x_{cd}).$$

We now have all the products of x, y of degree two in terms of the 135 products in (16) except the products of the type $x_{ac}y_{bc}$. If the value of y_{bc} given in (14.2) be multiplied by x_{ac}, and terms replacèd by using (17) and (18) we find that

$$\begin{aligned}(19)\qquad 4x_{ac}y_{bc}=3x^2_{ac}&+x^2_{bc}-x^2_{ba}+\Sigma_3(x^2_{ad}-x^2_{cd}+x^2_{ef}-x^2_{bd})+x_{ac}y_{ac}\\ &+x_{ac}\Sigma_3(y_{bd}+y_{ef})+y_{ac}\Sigma_3(x_{ef}-x_{bd})+(x_{bc}-x_{ba})\Sigma_3y_{ef}\\ +\Sigma_3\{(x_{ef}y_{ef}-x_{bd}y_{bd})&+(x_{ef}y_{bd}-x_{bd}y_{ef})+(x_{ad}-x_{cd})(y_{be}+y_{bf}+y_{ef})\},\end{aligned}$$

where Σ_3 indicates the cyclic advance of d, e, f.

Let (14.2) be multiplied by y_{ab} to yield

$$(20)\qquad 24y^2_{ab}=-8x_{ab}y_{ab}-8y_{ab}\Sigma_6x_{cd}+4y_{ab}\Sigma_8x_{ac}.$$

The first two terms on the right are in the aggregate (16). The third type can be expressed by using (19) in terms of this aggregate. When this is done there will occur on the right no terms in y^2_{ij}, whence the 15 relations (20) will serve as the 15 linearly independent relations connecting the 135 products.

In terms of the division $ab, cdef$ of the indices this modified relation (20) reads:

$$\begin{aligned}(21)\qquad 24y^2_{ab}=8x^2_{ab}&-4\Sigma_8x^2_{ac}+8\Sigma_6x^2_{cd}-8x_{ab}y_{ab}-2\Sigma_8x_{ac}y_{ac}\\ &+4\Sigma_6x_{cd}y_{cd}+4\Sigma_6x_{ab}y_{cd}-2\Sigma_{24}x_{ac}y_{bd}-2\Sigma_{24}x_{ac}y_{de}\\ &-8\Sigma_6x_{cd}y_{ab}+4\Sigma_{24}x_{cd}y_{ae}+4\Sigma_6x_{cd}y_{ef}.\end{aligned}$$

If this identity be expressed in terms of the Göpel squares by using (15) it reads:

$$\begin{aligned}(22)\qquad 3\Sigma_8x^2_{ac}&+4y^2_{ab}-2\Sigma_8y^2_{ac}-2\Sigma_6y^2_{cd}-4[ab]^2-\Sigma_8[ac]^2+2\Sigma_6[cd]^2\\ &+2\Sigma_6[ab,cd]^2-\Sigma_{24}[ac,bd]^2-\Sigma_{24}[ac,de]^2\\ &-4\Sigma_6[cd,ab]^2+2\Sigma_{24}[cd,ae]^2+2\Sigma_6[cd,ef]^2=0.\end{aligned}$$

Hence

(23) *The* 15 *relations* (22) *formed for divisions $ab, cdef$ of the indices constitute the* 15 *independent linear relations among the Göpel squares.*

By adding the three relations (22) formed for isolated ab, cd, ef, and similarly for isolated ab, ac, bc, by subtracting $6(\sigma_{x^2} - \sigma_{y^2}) = 0$ and deleting the factor 6 in each case, two sets of respectively 15 and 20 relations are obtained, namely:

$$(24.1) \quad \Sigma_{12}([ab, ce]^2 - [ce, ab]^2) - \Sigma_3(x^2_{ab} - y^2_{ab}) = 0 \quad \{\text{division } ab, cd, ef\};$$

$$(24.2) \quad \Sigma_9([ef, ad]^2 - [cd, ab]^2) - \Sigma_3(x^2_{de} - y^2_{ab} + [de]^2 - [ab]^2) = 0 \quad \{\text{division } abc, def\}.$$

From these 35 relations (24) the 15 independent relations (22) can be recovered as in [8, II, p. 21].

The Göpel squares appear as factors in products of the d_{ijk}. In fact the individual d_{ijk} have quite simple expressions in terms of these squares. For, the first cofactors in the two determinants (8.2) are, respectively

$$(x_{ac} + y_{ef})(x_{ac} + y_{de}) - (x_{bc} + y_{df})(x_{ab} + y_{df}),$$
$$(x_{ad} + y_{bf})(x_{ad} + y_{ce}) - (x_{ce} + y_{ad})(x_{bf} + y_{ad}).$$

If we apply (17) and (18) respectively to the pairs of terms,

$$x_{ef}y_{de} - x_{bc}x_{ab}, \qquad y_{bf}y_{ce} - x_{ce}x_{bf},$$

then only squared terms remain, and we have

$$(25.1) \quad 2^7 \cdot d_{123} = (x^2_{ab} + x^2_{ac} + x^2_{bc}) - (y^2_{de} + y^2_{df} + y^2_{ef}),$$

$$(25.2) \quad 2^7 \cdot d_{567} = (x^2_{ad} + x^2_{ce} + x^2_{bf}) - (y^2_{ad} + y^2_{ce} + y^2_{bf}).$$

It is then clear that

(26) *The* 56 *linear five-term relations which connect the* 35 d_{ijk} *are reduced, by virtue of the definitions* (25), *to the single linear relation,* $\sigma_{x^2} - \sigma_{y^2} = 0$, *among the Göpel squares.*

When we pass to one of the other 35 self-associated P_8^3's defined by a ternary quartic, this single linear relation, connecting two sets of 15 Göpel squares, is replaced by one of the 35 relations (24). Thus it is relatively simple to define as in (25) any one of the 630 [cf. text after 8.2] linear 8-point invariants attached to the ternary quartic in terms of six properly chosen Göpel squares.

We consider now the matrix,

$$(27) \qquad M_{567} = \left\| \begin{matrix} d_{125}\, d_{345} & d_{126}\, d_{346} & d_{127}\, d_{347} \\ d_{135}\, d_{245} & d_{136}\, d_{246} & d_{137}\, d_{247} \\ d_{145}\, d_{235} & d_{146}\, d_{236} & d_{147}\, d_{237} \end{matrix} \right\| .$$

The elements in the first row are the products which occur in the expression for the invariant $\delta_{12,34}$ [cf. **12** (4)]; those in the first column are similarly connected with the invariant $\delta_{58,67}$. The 35 matrices M_{567} thus bring in the 35. 6 = 210 invariants $\delta_{ij,kl}$. If the elements of the matrix M_{567} are each multiplied by d_{567} to produce the matrix $d_{567}M_{567}$, the resulting elements can be expressed by using (6) and (8.1) in the form:

$$(28) \qquad \begin{aligned} d_{127}d_{347}d_{567} &= \Delta_1 \cdot\cdot\cdot \Delta_7 \cdot \Delta_7{}^2(127)(347)(567) \\ &= D \cdot [ad]^2, \\ D &= \Delta_1\Delta_2 \cdot\cdot\cdot \Delta_7. \end{aligned}$$

Hence, making use of the irrational form [cf. **12** (9)] of $\delta_{ij,kl}$:

(29) *The elements of the matrix* $D^{-1} \cdot d_{567} \cdot M_{567}$ *are the squares of the elements in the second determinant of* (8.2). *Thus the* 210 *irrational conditions for the self-association of* $P_8{}^3$ *appear as linear three-term relations connecting the Göpel invariants, six of these conditions being obtained from the six lines of each of the* 35 *matrices* M_{567}.

Let

$$(30) \qquad \begin{aligned} f(x;\, y) &= f(x_1, x_2, x_3;\, y_1, y_2, y_3) \\ &= x_1y_1 + x_2y_2 + x_3y_3 - x_1(y_2 + y_3) - x_2(y_3 + y_1) - x_3(y_1 + y_2) \end{aligned}$$

be the polarized form of the quadratic expression **12** (9). Thus the conditions that $P_8{}^3$ be self-associated are $f(x;x) = 0$ where x_1, x_2, x_3 are the elements of any line of the 35 matrices M_{567}. We wish to prove that

(31) *The expression* $f(x;y)$ *formed for any two parallel lines* $x_1, x_2, x_3;$ y_1, y_2, y_3 *of any one of the* 35 *matrices* M_{567} *has the fixed value* $-2D^2$.

We begin with the determinants,

$$(32) \qquad \begin{aligned} D_{8,7} &= \left| \begin{matrix} (3418)(5618) & (5318)(4618) \\ (3428)(5628) & (5328)(4628) \end{matrix} \right| , \\ D_{7,8} &= \left| \begin{matrix} (5627)(3427) & (4627)(5327) \\ (5617)(3417) & (4617)(5317) \end{matrix} \right| . \end{aligned}$$

$D_{8,7} = 0$ is the condition that there be a quadric cone with node at p_8 and on all of the other points p except p_7. When $P_8{}^3$ is self-associated, the condition $D_{8,7} = 0$ implies the condition $D_{7,8} = 0$, and $P_8{}^3$ is a set of eight points on a cubic curve. Indeed the like placed determinants $(ijkl)$ in $D_{8,7}$

and $D_{7,8}$ are complementary and therefore proportional. In the product $D_{8,7}D_{7,8}$ two of the four terms are products of four d_{ijk}'s, and the other two are products of two invariants of the second degree [cf. **12** (1)],

$$D_{8,7}D_{7,8} = d_{134}d_{156}d_{235}d_{246} + d_{234}d_{256}d_{135}d_{146}$$
$$- [18, 27;\ 36, 45]\ [17, 28;\ 36, 45] - [28, 17;\ 36, 45]\ [27, 18;\ 36, 45].$$

If the last two terms are expressed in terms of the d_{ijk} from **12** (5) with $\delta_{ij,kl} = 0$, we have

$$(33) \qquad -2D_{8,7}D_{7,8} = f(d_{134}d_{156},\ d_{135}d_{146},\ d_{136}d_{145};\ d_{234}d_{256},\ d_{235}d_{246},\ d_{236}d_{245}).$$

The arguments of $f(x; y)$ as here obtained are found in two parallel rows of the matrix M_{127}. Since however the product $D_{8,7}D_{7,8}$ is invariant under permutation of the points $p_1, \cdots, p_6$ the same value is obtained from any one of the 15 matrices M_{ij7} $(i, j = 1, \cdots, 6)$. It is also clear that, if the transpositions (17) and (28) be applied to (32), the same expression f is obtained on the right in (33). Thus the 35. 6 expressions f obtained from two parallel lines of the 35 matrices M_{567} give rise to 15 forms for each of the 28 $D_{i,j}D_{j,i}$ $(i, j = 1, \cdots, 8)$, two of the latter products being represented by the same expression f. In order to evaluate these expressions f, denote in conformity with (29) two parallel rows of $D^{-1} \cdot d_{567} \cdot M_{567}$ by r_1^2, r_2^2, r_3^2; s_1^2, s_2^2, s_3^2 where $r_1 + r_2 + r_3 = 0$ and $s_1 + s_2 + s_3 = 0$. Then on eliminating r_3, s_3 from $f(r_1^2, r_2^2, r_3^2; s_1^2, s_2^2, s_3^2)$ we find the value $-2(r_1s_2 - r_2s_1)^2$. But this, according to (8. 2), is $-2d^2_{567}$. Hence $f(x; y)$, formed for two parallel rows of M_{567}, is precisely $-2D^2$ and (31) is proved. Also we see that

$$(34) \qquad D_{i,j}D_{j,i} = D^2 = \Delta_1^2\Delta_2^2 \cdots \Delta_7^2 \qquad (i, j = 1, \cdots, 8).$$

The relations (34) are found in ([3], pp. 177-178) as consequences of a certain system of equations [cf. [3], pp. 75-76] in which the values of certain constants attached to the 64 odd and even theta functions $(p = 3)$ of the first order are related to the 63 discriminant conditions and to a factor of proportionality r. It there appears that

$$(35) \qquad D = r\vartheta^4(0).$$

Thus the vanishing of D is the condition that the hyperelliptic case appear, in which P_8^3 is a set of eight points on a twisted cubic N^3. Taking account of (31) we see that

(36) *The self-associated P_8^3 is the hyperelliptic P_8^3 on a norm-curve N^3 if any one of the expressions $f(x; y)$ of* (31) *vanishes.*

It must be emphasized that if $D = 0$, an additional quartic relation

satisfied by the d_{ijk}'s, the present representation of the d_{ijk}'s in terms of the Göpel invariants attached to Q_7^2 entirely fails. For, the functions being hyperelliptic, they cannot be attached to a normal quartic envelope with nodes at Q_7^2. The normal algebraic form is rather the binary octavic, and the Göpel invariants for this case have been defined anew, and quite differently, in **8**.

The irrational invariants $D_{i,j}$ appear (under the name of Pascalians) in a memoir of H. S. White [12] who finds that the square, $(8\xi)^2$, of the point p_8 of the self-associated P_8^3 is

$$(37) \qquad (1\xi)^2/D_{1,8} + (2\xi)^2/D_{2,8} + \cdots + (7\xi)^2/D_{7,8} = 0.$$

This result may be obtained readily, and in more symmetric form, by the following use of the discriminant conditions. Let $\lambda_1(1\xi)^2 + \cdots + \lambda_8(8\xi)^2 = 0$ be the identity which expresses that quadrics on $p_1, \cdots, p_7$ all pass through p_8. Polarize this as to η, and set $(\xi x) = (123x)$, $(\eta x) = (456x)$. Then $(1237)(4567)\lambda_7 + (1238)(4568)\lambda_8 = 0$. Setting

$$(1237) = \epsilon_{12}\epsilon_{13}\epsilon_{17}\epsilon_{23}\epsilon_{27}\epsilon_{37}\epsilon_{1237}[\epsilon_{1237} = \epsilon_{4568}], \text{ etc.};$$

taking account also of the fact that, for the complementary determinants, $(1237) = -(4568)\cdot\kappa$, and $(1238) = (4567)\cdot\kappa$, we find that $E_7\lambda_7 = E_8\lambda_8$, where

$$(38) \qquad E_i = \epsilon_{ij}\epsilon_{ik}\cdots\epsilon_{ip} \qquad (i, j, \cdots, p = 1, \cdots, 8).$$

Hence the above identity reads as follows:

$$(39) \qquad (1\xi)^2/E_1 + (2\xi)^2/E_2 + \cdots + (8\xi)^2/E_8 \equiv 0.$$

If we multiply this by E_8, and take account of

$$(40) \qquad D_{i,j}/D = D/D_{j,i} = E_i/E_j \qquad [\text{cf. }^3\text{, p. 179 (13)}],$$

the left member of (37) becomes $-D\cdot(8\xi)^2$, which is White's result. The relation (39) has the advantage over (37) in symmetry. It is to be observed however that the coefficients $D_{i,j}$ in (37) can be expressed rationally in terms of the coördinates of P_8^3 while the coefficients E_i in (39) cannot be so expressed, since they depend upon an irrational factorization of $D_{i,j}$.

The quantity, $D = \Delta_1\Delta_2\cdots\Delta_7$, whose vanishing indicates the hyperelliptic case, appears in another connection. If the equation of the quartic envelope with nodes at Q_7^2 be written in the form, $f(\xi^4, q^{10}) = 0$ [cf. [3], pp. 191-192], an invariant of the envelope of degree $3l$, and therefore of degree $30l$ in each point q, contains the factor D^{2l}, and a further significant factor of degree $2l$ in the Göpel invariants. Thus the non-vanishing of D is necessary for the representation in terms of Q_7^2. When D is zero, many of the formulae

just derived require revision. For example, according to (28), $d_{127}d_{347}d_{567}$ is zero if $D=0$, which is evidently an absurdity for eight points on N^3. The proper formula in that case is **14** (2.1).

In order to obtain an expression for D in terms of the d_{ijk} for the self-associated P_8^3 we take the particular case of **11** (3), (5) when $\delta_{12,34}=0$, namely

$$\begin{aligned}2[12,34;\ 56,78] &= d_{125}d_{345}+d_{126}d_{346}-d_{127}d_{347}\\ &= \Delta_1\Delta_2\Delta_3\Delta_4\{\Delta_5^2(125)^2(345)^2+\Delta_6^2(126)^2(346)^2-\Delta_7^2(127)^2(347)^2\}.\end{aligned}$$

By virtue of the linear identity (5) this becomes

$$(41)\quad \begin{aligned}[12,34;\ 56,78] &= -\Delta_1\Delta_2\cdots\Delta_6\cdot(125)(346)\cdot(126)(345)\\ &= (d_{125}d_{345}+d_{126}d_{346}-d_{127}d_{347})/2.\end{aligned}$$

We now examine a two-row determinant of the matrix M_{567}, which yields by virtue of (6)

$$(42)\quad Q=-\begin{vmatrix} d_{145}d_{235} & d_{146}d_{236}\\ d_{135}d_{245} & d_{136}d_{246}\end{vmatrix}=\Delta_1^2\cdots\Delta_6^2\begin{vmatrix}(135)^2(245)^2 & (136)^2(246)^2\\ (145)^2(235)^2 & (146)^2(236)^2\end{vmatrix}.$$

The resulting determinant is an alternating invariant of $Q_6^2=q_1,\cdots,q_6$, which (cf. [1], I, §§ 4, 5) has the value

$$Q=\dot{-}\,\Delta_1^2\cdots\Delta_6^2\Delta_7\{(135)(245)(146)(236)+(136)(246)(145)(235)\}.$$

The brace is an invariant of the second degree of Q_6^2 denoted by $\overline{ad}$ in [[1], pp. 172-173 (41), (46)]. It can be expressed as of degree two in the linear invariants of Q_6^2 (products such as occur in the second expression of (41) above) as follows [cf. [1], I, p. 175]:

$$\begin{aligned}\overline{ad}=\{&(351)(462)\cdot(352)(461)\\ &+(361)(452)\cdot(451)(362)-(561)(342)\cdot(562)(341)\}.\end{aligned}$$

Hence, applying (41), we find that

$$(43)\quad \begin{aligned}Q &= D\{[35,46;\ 12,78]+[36,45;\ 12,78]-[34,56;\ 12,78]\}\\ &= -D\cdot\Delta_1\cdots\Delta_6\,\overline{ad}\\ &= 1/2\,D\cdot\{d_{351}d_{461}+d_{352}d_{462}-d_{357}d_{467}+d_{361}d_{451}\\ &\qquad +d_{362}d_{452}-d_{367}d_{457}-d_{341}d_{561}-d_{342}d_{562}+d_{347}d_{567}\}.\end{aligned}$$

Since, according to (43), the quantity $-2\Delta_1\cdots\Delta_6\,\overline{ad}$ is of degree two in the d_{ijk}, and since [cf. [1], I, p. 173 (44)]

$$(44)\qquad -d_2^2=\Delta_7^2=\overline{ad}\,\overline{be}+\overline{be}\,\overline{cf}+\overline{cf}\,\overline{ad},$$

there follows that

(45) *The quantity* $D^2 = \Delta_1^2 \cdots \Delta_7^2$ *is rationally and integrally expressible in terms of the* d_{ijk} *of degree four, and in terms of the Göpel invariants of degree eight; the quantity* D *however is only rationally expressible in terms of the* d_{ijk} *and the Göpel invariants as in* (43).

When the functions are hyperelliptic the two-row determinants of M_{567}, such as Q, all vanish [cf. **14** (3)]. If as before two rows of M_{567} are x_1, x_2, x_3 and y_1, y_2, y_3, then $f(x;x) = 0$ and $f(y;y) = 0$, since P_8^3 is self-associated. Also $f(x;y) = 0$ since the functions are hyperelliptic [cf. (36)]. Since $f(x;x)$ is a non-degenerate quadratic form, these three conditions require that $x_i = \lambda y_i$ $(i = 1, 2, 3)$, whence the determinants formed from the two rows vanish. All of these conditions are of degree four in the d_{ijk}, but the conditions given by the vanishing of the determinants are not linearly expressible in terms of the others. In fact if we write,

$$f(x;\, x) = (x_3 - x_1 - x_2)^2 - 4x_1x_2, \qquad f(y;\, y) = (y_3 - y_1 - y_2)^2 - 4y_1y_2,$$
$$f(x;\, y) = (x_3 - x_1 - x_2)(y_3 - y_1 - y_2) - 2(x_1y_2 + x_2y_1),$$

then

$$\begin{aligned}(46) \quad \{2(x_1y_2 - x_2y_1)\}^2 &= f(x;\, x)f(y;\, y) - f^2(x;\, y)\\ &+ 4\{x_1x_2\, f(y;\, y) + y_1y_2\, f(x;\, x) - (x_1y_2 + x_2y_1)f(x;\, y)\}.\end{aligned}$$

Thus the squares of the determinants are in the *modulus* determined by the forms $f(x;x)$, $f(y;y)$, $f(x;y)$, though the determinants themselves are not.

We complete the set of formulae (28) as follows [cf. (8. 1)]:

$$(47) \quad \begin{aligned} d_{127}d_{347}d_{567} &= D \cdot [ad]^2,\\ d_{126}d_{346}d_{567} &= D \cdot [ce, bf]^2,\\ d_{135}d_{146}d_{234}d_{256}d_{127}d_{367}d_{457} &= D^3 \cdot [cf +]^2.\end{aligned}$$

Of the first type there are 15; of the second, 90, and of the third, 30. Since, according to (45), D^2 is rational, integral, and of degree four in the d_{ijk}, then any Göpel square multiplied by D^3 is rational, integral, and of degree 7 in the d_{ijk}. Since, as remarked above, any invariant of degree $3l$ of the ternary quartic attached to Q_7^2 has an effective factor of degree $2l$ in the Göpel squares, there follows that

(48) *The invariants of the ternary quartic of degree* $3l$ *can be expressed as polynomials in the* d_{ijk} *of degree* $7l$.

Even for the generic self-associated P_8^3 the d_{ijk} satisfy two systems of quartic relations. The first system, sufficient alone to define the self-associated P_8^3, arises from the rows x_i, y_i, z_i, and the columns of the matrices M_{567}. These are expressed by

$$f(x;x)=0, f(y;y)=0, f(z;z)=0, f(1;1)=0, f(2;2)=0, f(3;3)=0,$$

where $f(i;i)=(z_i-y_i-x_i)^2-4x_iy_i$ $(i=1,2,3)$. The second system results from (31). For, $f(x;y)=f(x;z)=-2D^2$ yields the identity, $f(x;y-z)=0$, also of degree four in the d_{ijk}. This second system is algebraically, but apparently not linearly, dependent upon the first system. For, from $f(i;i)=0$ we get $z_i-y_i=x_i+2(x_iy_i)^{1/2}$. Hence

$$\begin{aligned} f(x;z-y) &= f(x;x)+2f[x;(xy)^{1/2}] \\ &= 2f[x;(xy)^{1/2}]=2\Sigma_i(x_iy_i)^{1/2}\,(x_i-x_j-x_k). \end{aligned}$$

But from $f(x;x)=0$ there follows $x_i-x_j-x_k=2(x_jx_k)^{1/2}$. Hence $f(x;z-y)=2(x_1x_2x_3)^{1/2}\,[\Sigma_i(y_i)^{1/2}]$. But $\Sigma_i(y_i)^{1/2}=0$ since $f(y;y)=0$, whence $f(x;z-y)=0$. It is still possible of course that this system of quartic relations, $f(x;y)-f(x';y')=0$, where x,y are two parallel lines of one matrix and x',y' two of another matrix, may be linearly dependent on the entire system $f(x'';x'')=0$. We merely have found nothing to indicate such linear dependence.

In the hyperelliptic case the d_{ijk} satisfy a system of 14 cubic relations which may be expressed more conveniently in the form of 35 relations,*

$$(49)\qquad \begin{aligned} R_{123} &= \Sigma\,\epsilon_{ijkl}\,d_{ij1}\,d_{ik2}\,d_{il3}=0 \qquad (i,j,k,l=4,5,6,7);\\ R_{567} &= \Sigma\,\epsilon_{ijkl}\,d_{ij5}\,d_{ik6}\,d_{il7}=0 \qquad (i,j,k,l=1,2,3,4), \end{aligned}$$

where ϵ_{ijkl} is the sign of the permutation $ijkl$ from the natural order. These relations do not exist in the present case but it is a matter of some interest to find the values which the right members take.

R_{567} is associated with the division, ad, bf, ce, of the indices, and it is

$$(50.1)\qquad \begin{aligned} R_{567} &= D\{\Sigma_{12}[bf,ca]^2-\Sigma_{12}[ca,bf]^2\}\\ &= D\{\Sigma_3(x^2{}_{bf}-y^2{}_{bf})\} \qquad [\text{cf. } (24.1)]\\ &= 2^7\cdot D\cdot d_{567} \qquad [\text{cf. } (25.2)]. \end{aligned}$$

R_{123} is associated with the ordered division, abc, def, of indices, and it is

$$(50.2)\qquad \begin{aligned} R_{123} &= D\{\Sigma_9[ab,cd]^2-\Sigma_9[af,de]^2+\Sigma_3[ab]^2-\Sigma_3[de]^2\}\\ &= D\{\Sigma_3x^2{}_{ab}-\Sigma_3y^2{}_{de}\} \qquad [\text{cf. } (24.2)]\\ &= 2^7\cdot D\cdot d_{123} \qquad [\text{cf. } (25.1)]. \end{aligned}$$

That $R_{567}=0$ in the hyperelliptic case is a consequence of the fact that the

* These relations are given correctly in II **10** (20). In II **10** (21) however ϵ_{ijkl} was incorrectly omitted because it was not observed that the change of sign which occurs in $\Delta^{1/2}$ is counteracted by a change of sign in the d_{ijk} as defined earlier in **10** (1).

Göpel invariants for this case satisfy quadratic relations. Indeed in this case we have [cf. **9** (3), (5), (6)]

$$\begin{vmatrix} \sigma_3+\bar\sigma_2 & \sigma_2+\bar\sigma_1 \\ \sigma_2+\bar\sigma_3 & \sigma_1+\bar\sigma_2 \end{vmatrix} = \begin{vmatrix} [bc,cf] & [cf,ad] \\ [cf,be] & [ad,cf] \end{vmatrix} = 0.$$

In the more general case however this determinant has the value $-64d_{127}$ [cf. (8.2)].

14. The hyperelliptic P_8^3 on the cubic space curve N^3. We first examine the formulae developed above for more general sets P_8^3 in order to see in what form they persist in the present case. Since this special P_8^3 on N^3 is itself a self-associated set, the quintic relations of **12** (35) are satisfied because the quartic relations **12** (38) [the relations $f(x;x)=0$ of **13** (31) are satisfied. Fundamental differences in the hyperelliptic and the more general case are as follows. The hyperelliptic Göpel invariants are products of *four* discriminant factors of the underlying octavic instead of *seven* discriminant factors of the underlying quartic curve. The two sets of invariants satisfy the same linear relations except that in the hyperelliptic case we have an additional linear relation, $\sigma_x=0$, $\sigma_y=0$, replacing $\sigma_x+\sigma_y=0$ in **13** (12). The additional linear relation entails the existence of 14 additional quadratic relations [and also $(\sigma_x)^2=0$], which can be used more conveniently in the form of 35 relations given in **9** (7).

In the hyperelliptic case the d_{ijk} of P_8^3 are of degree three in the Göpel invariants rather than the degree *two* [cf. **13** (8.2)]. In fact [cf. **10** (7); **9** (3), (5)] the hyperelliptic values are

$$(1)\qquad \begin{aligned} d_{127} &= [ad]\cdot[be]\cdot[cf], \\ d_{123} &= [ab,de]\cdot[ac,ef]\cdot[bc,df]. \end{aligned}$$

The quantity which now is supposed not zero is the discriminant Δ of the octavic rather than the quantity D^2 of **13** (34). Though D is rationally, but not integrally, expressible in terms of the Göpel invariants, yet both Δ and $\Delta^{1/2}$ are expressible rationally and integrally in terms of the hyperelliptic Göpel invariants [cf. **10** (15) and [6]].

The formulae **13** (47) now take the simpler form

$$(2.1)\qquad d_{127}d_{347}d_{567}=\Delta^{1/2}\cdot(12)^2(34)^2(56)^2(78)^2,$$

$$(2.2)\qquad d_{135}d_{146}d_{234}d_{256}d_{127}d_{367}d_{457}=\Delta^{3/2}.$$

These enable us to prove that

(3) *In the hyperelliptic case the* 35 *matrices* M_{567} *of* **13** (27) *have the rank one.*

For, the cofactor of an element of this matrix, multiplied by the element and by d_{567}, yields two terms of type (2.2) with opposite signs.

We observe also that, according to (2.1), the matrix $\Delta^{-1/2}\cdot d_{567}\cdot M_{567}$ has elements $a_{ij}=\alpha_i\beta_j$, where

$$(4)\qquad \begin{array}{lll} \alpha_1=(12)^2(34)^2, & \alpha_2=(13)^2(42)^2, & \alpha_3=(14)^2(23)^2; \\ \beta_1=(58)^2(67)^2, & \beta_2=(68)^2(75)^2, & \beta_3=(78)^2(56)^2. \end{array}$$

From this again the irrational conditions for self-association are versions of the three-term relations among the hyperelliptic Göpel relations. Also the theorem (3), and the vanishing of all the expressions $f(x;y)=D^2$ [cf. **13** (31)] become self-evident.

We again call attention to the cubic relations connecting the d_{ijk} give in **13** (49) as $R_{123}=0$, $R_{567}=0$, which in the non-hyperelliptic case ta the forms given in **13** (50.1), (50.2). The satisfaction of these cub relations alone does not seem sufficient to characterize the d_{ijk} as belonging to a self-associated P_8^3. For, if these relations are satisfied, the individual terms are proportional to squares of Göpel invariants [cf. (2.1)]. Thus the values of the Göpel invariants are determined to within sign, and the quadratic relations satisfied by the Göpel invariants [cf. **9** (7), (22)] necessaril hold. But also the linear three-term relations among the Göpel invariants must be satisfied, and these, as has just been pointed out, are satisfied by virtue of the fourth degree conditions for self-association.

We obtain finally the rational integral expressions of degree five in the d_{ijk} which are equal to the linear Göpel invariants each multiplied by Δ. Denote by $(i_1i_2\cdots;\ j_1j_2\cdots)$ the product $\Pi(i_rj_s)$ of differences of roots of the octavic. Then the following identity is obvious:

$$\Delta^{1/2}/(12)(34)(56)(78)=\Delta^{1/2}(12;\ 34)(56;\ 78)/d_{567}.$$

In terms of the quantities introduced in (4) we have

$$\begin{gathered} -2(12;\ 34)=\alpha_1-\alpha_2-\alpha_3;\qquad -2(56;\ 78)=\beta_3-\beta_1-\beta_2; \\ 4\Delta^{1/2}/(12)(34)(56)(78)=\Delta^{1/2}(\alpha_1-\alpha_2-\alpha_3)(\beta_3-\beta_1-\beta_2)/d_{567}. \end{gathered}$$

If the product on the right is evaluated with reference to (4) and (2.1) we have

$$(5.1)\qquad 4\Delta^{1/2}/(12)(34)(56)(78)=K_{12,34,56,78},$$

$$(5.2)\qquad \begin{aligned} K_{12,34,56,78}=&\ d_{127}d_{347}+d_{135}d_{245}+d_{136}d_{246}+d_{145}d_{235}+d_{146}d_{236} \\ &-[d_{125}d_{345}+d_{126}d_{346}+d_{137}d_{247}+d_{147}d_{237}]. \end{aligned}$$

The negative terms in $K_{12,34,56,78}$ are those elements of M_{567} not in a line with

$d_{127}d_{347}$, the positive terms are the remaining elements of M_{567}. On multiplying this by (2.1) we have

$$(6) \qquad 4\Delta \cdot (12)(34)(56)(78) = d_{127}d_{347}d_{567}K_{12,34,56,78},$$

the expression on the right being rational, integral, and of degree five in the d_{ijk}.

In II **10** (7), (8) the values of the invariants (B) of the underlying octavic, the d_{ijk}, are given as polynomials of the third degree in the invariants (A), the Göpel invariants. The formula (6), which furnishes the values of the invariants (A), each multiplied by the discriminant, as polynomials of the fifth degree in the invariants (B), represents the inverse Tschirnhaus transformation between the linear system of irrational invariants (A) and the linear system of irrational invariants (B).

This is the most favorable form the inverse transformation can take. Indeed, if, for generic p,

$$(\Delta^{1/2})^k \cdot A = P^l(B)$$

where $P^l(B)$ is a polynomial of order l in the invariants B, then, by equating the weights on both sides, we get

$$(p+1)(2p+1)k + (p+1) = l(p+1)p, \quad \text{i.e.,} \quad p(l-2k) = k+1.$$

Hence p must divide $k+1$ and the most favorable values of k and l are $k = p-1$, $l = 2p-1$, which are the values in (6).

REFERENCES.

[1] A. B. Coble, "Point sets and allied Cremona groups," *Transactions of the American Mathematical Society*, I, Vol. 16 (1915), pp. 155-198; II, Vol. 17 (1916), pp. 345-385.

[3] A. B. Coble, "Algebraic geometry and theta functions," *Colloquium Publications of the American Mathematical Society*, New York, Vol. 10 (1929).

[6] C. M. Huber, "On complete systems of irrational invariants of associated point sets," *American Journal of Mathematics*, Vol. 49 (1927), pp. 251-267.

[8] A. B. Coble, "Hyperelliptic functions and irrational binary invariants," *American Journal of Mathematics*, I, Vol. 54 (1932), pp. 425-452; II, Vol. 55 (1933), pp. 1-21.

[12] H. S. White, "The associated point of seven points in space," *Annals of Mathematics*, Ser. 2, Vol. 23 (1923), pp. 301-306.

ON A CERTAIN LINEAR ∞^r-SYSTEM OF r-IC HYPERSURFACES IN r-SPACE.*

By B. C. Wong.

1. *Introduction.* Let $r+1$ general $(r-2)$-spaces $S_{r-2}^{(j)}$ $[j=0,1,\cdots,r]$ be given in an r-space S_r. For a hypersurface $V^r{}_{r-1}$ of order r to pass through these $r+1$ given $(r-2)$-spaces is equivalent to

$$N=\sum_{i=1}^{t}(-1)^{i-1}\binom{r+1}{i}\binom{2r-2i}{r}$$

linear conditions where $t=r/2$ if r is even and $t=(r-1)/2$ if r is odd. Since it takes $\binom{2r}{r}-1$ simple conditions to determine a $V^r{}_{r-1}$ in S_r, the dimension of the linear system, $|V|$, of the $V^r{}_{r-1}$'s passing through the $r+1$ given $(r-2)$-spaces is

$$\binom{2r}{r}-1-N=\sum_{i=0}^{t}(-1)^{i}\binom{r+1}{i}\binom{2r-2i}{r}-1=r.$$

This ∞^r-system, $|V|$, of r-ic hypersurfaces has already been discussed in a brief note by Veneroni † but the value of this note has been greatly diminished because of a serious error which the author made in the reasoning concerning the variety V_{r-2} common to all the members of the system. It is our present purpose to rectify this error and incidentally present a few details not given by Veneroni. We shall then derive the equation of a general $V^r{}_{r-1}$ of the system and thereby establish a birational r-ic transformation ‡ between S_r and another r-space R_r. Finally we shall obtain the conditions for which the transformation be involutorial within S_r and shall relate this involutorial transformation as a very special case to a certain general type of r-ic involutorial transformations in S_r.

2. *Veneroni's error.* According to Veneroni all the $V^r{}_{r-1}$'s of $|V|$ have in common an $(r-2)$-dimensional variety $V_{r-2}^{(r-1)(r-2)}$ of order $(r-1)(r-2)$ and this variety is the locus of the ∞^{r-3} lines incident with all the $r+1$ given $(r-2)$-spaces $S_{r-2}^{(j)}$ and is the residual intersection of any two of the $r+1$ $(r-1)$-ic ruled hypersurfaces V_{r-1}^{r-1}'s each passing through r of the $r+1$ given $(r-2)$-spaces.§ These conclusions are erroneous.

* Presented to the American Mathematical Society, June 21, 1929.

† "Sopra una trasformazione birazionale fra due S_n," *Rendiconti Ist. Lomb.* (2), Vol. 34 (1901), pp. 640-644.

‡ This transformation is the subject of Veneroni's note just referred to.

§ Segre reported these results, uncorrected, in a footnote in his "Mehrdimensionale Räume," *Encyklopädie der Mathematischen Wissenschaften*, III$_2$, 7, p. 967.

In the first place, the locus of the ∞^{r-3} lines incident with $r+1$ general S_{r-2}'s in S_r is an $(r-2)$-dimensional variety which is not of order $(r-1)(r-2)$ but of order $(r+1)(r-2)/2$. We denote this variety by $M_{r-2}^{(r+1)(r-2)/2}$. It is this $M_{r-2}^{(r+1)(r-2)/2}$ that is common to all the $V^r{}_{r-1}$'s of $|V|$.* In the second place, it is true that the residual intersection of any two of the V_{r-1}^{r-1}'s each passing through r of the $r+1$ given $(r-2)$-spaces is a $V_{r-2}^{(r-1)(r-2)}$. But this $V_{r-2}^{(r-1)(r-2)}$ is composed of two varieties one of which is the $M_{r-2}^{(r+1)(r-2)/2}$ above and the other is a variety $V_{r-2}^{(r-2)(r-3)/2}$ of order $(r-2)(r-3)/2$. This latter variety, $V_{r-2}^{(r-2)(r-3)/2}$, is not contained in all the $V^r{}_{r-1}$'s of the system and it is the locus of ∞^{r-4} planes each meeting the $r-1$ $(r-2)$-spaces common to the two V_{r-1}^{r-1}'s in lines and the two remaining $(r-2)$-spaces each in a point. In each of the planes of this locus there is only one line incident with all the $r+1$ given $(r-2)$-spaces and it belongs to the other variety, $M_{r-2}^{(r+1)(r-2)/2}$.

Let us illustrate. For $r=4$, we have ∞^4 $V_3{}^4$'s passing through five given planes $\alpha^{(j)}$ $[j=0,1,\cdots,4]$ in S_4. The locus of the ∞^1 lines incident with these five planes is a ruled quintic surface F^5. The same five planes determine five ruled $V_3{}^3$'s. Let $W^{(0)}$ be the $V_3{}^3$ with its rulings incident with $\alpha^{(1)}$, $\alpha^{(2)}$, $\alpha^{(3)}$, $\alpha^{(4)}$ and $W^{(1)}$ be the one with its rulings incident with $\alpha^{(0)}$, $\alpha^{(2)}$, $\alpha^{(3)}$, $\alpha^{(4)}$. Now $W^{(0)}$ and $W^{(1)}$, having already the three planes $\alpha^{(2)}$, $\alpha^{(3)}$, $\alpha^{(4)}$ in common, intersect in a sextic surface which is composed of F^5 and a plane $\beta^{(234)}$ determined by the three points $P^{(23)}\equiv\alpha^{(2)}\alpha^{(3)}$, $P^{(34)}\equiv\alpha^{(3)}\alpha^{(4)}$, $P^{(42)}\equiv\alpha^{(4)}\alpha^{(2)}$. This plane $\beta^{(234)}$ meets $\alpha^{(2)}$, $\alpha^{(3)}$, $\alpha^{(4)}$ in the lines $P^{(23)}P^{(42)}$, $P^{(23)}P^{(34)}$, $P^{(34)}P^{(42)}$, respectively, and meets $\alpha^{(0)}$ in a point A and $\alpha^{(1)}$ in a point A'. The line AA' is the only line in this plane that is incident with the five given planes, and is therefore a ruling of F^5.

Having definitely established that the variety contained in all the $V^r{}_{r-1}$'s of $|V|$ is of order $(r+1)(r-2)/2$, we find that any two general members of $|V|$ have for residual intersection a $V_{r-2}^{r(r-1)/2}$. This variety will be regarded as the free intersection of the two hypersurfaces. In general, as can be easily verified or can be established by means of the transformation to be given subsequently, q members of $|V|$ intersect in a free $V^n{}_{r-q}$ of order $n=\binom{r}{r-q}$. Thus, r members have one free point of intersection, and $r-1$ members intersect in a normal rational curve C^r, etc.

---

* It is known that the locus of the ∞^{2r-k-2} lines that meet k given $(r-2)$-spaces in r-space is a V_{2r-k-1} of order $k!(2r-k-1)/r!(k-r+1)$. Putting $k=r+1$, we have the result above. See B. C. Wong, "On the loci of the lines incident with k $(r-2)$-spaces in S_r," *Bulletin of the American Mathematical Society*, Vol. 34 (1928), pp. 715-717.

It is to be noticed that the ruled variety $M_{r-2}^{(r+1)(r-2)/2}$ common to all the $V^r{}_{r-1}$'s of $|V|$ is such that the curve in which a general S_3 of S_r intersects it has $r(3r^3-14r^2+9r+26)/24$ apparent double points, and that it meets each of the given $(r-2)$-spaces $S_{r-2}^{(j)}$ in a V_{r-3}^{r-1}. The free intersection $V_{r-2}^{r(r-1)/2}$ of any two members of $|V|$ is met by a general S_3 in a curve with $r(r-1)(r-2)(3r-5)/24$ apparent double points. The free intersection $F^{r(r-1)/2}$ of $r-2$ members is met by a general S_{r-1} in a curve of deficiency $(r-1)(r-2)/2$.

It is to be noticed also that the given $S_{r-2}^{(j)}$ intersect t by t in $\binom{r+1}{t}$ $(r-2t)$-spaces $S_{r-2t}^{(k)}$ $[t=1,2,\cdots,r/2$ if r is even, $t=1,2,\cdots,(r-1)/2$ if r is odd; $k=1,2,\cdots,\binom{r+1}{t}]$. All the hypersurfaces of $|V|$ contain $S_{r-2t}^{(k)}$ t-ply but each hypersurface has other multiple varieties besides $S_{r-2t}^{(k)}$. Thus, for $r=4$, all the ∞^4 $V_3{}^4$'s passing through five given planes in S_4 have 10 common double points which are the intersections of the five given planes two by two. Each of these $V_3{}^4$'s has 10 other double points, two in each of the given planes. For $r=5$, all the $V_4{}^5$'s containing six given S_3's in S_5 contain doubly the 15 lines of intersection of the six S_3's two by two, but each $V_4{}^5$ has six double quintic curves each lying in one of the given S_3's. In general, each $V^r{}_{r-1}$ of $|V|$ has $r+1$ double varieties V_{r-4}'s of order $r(r-1)/2$ each lying in one of the $r+1$ given $(r-2)$-spaces. These results can best be derived by means of the transformation we are about to set up.

2. *The equation of $V^r{}_{r-1}$.* Let the $r+1$ given $(r-2)$-spaces be represented, without loss of generality, by the equations

$$S_{r-2}^{(j)}:\quad x_j=0,\qquad \sum_{i=0}^{r} a_i{}^{(j)}x_i=0,\qquad [i\neq j].$$

In order to derive the equation of a general $V^r{}_{r-1}$ of $|V|$, we find it convenient to write first the equations of the $r+1$ ruled V_{r-1}^{r-1}'s each passing through r of the $r+1$ given $(r-2)$-spaces. Let $W^{(j)}$ denote the V_{r-1}^{r-1} whose rulings are not incident with $S_{r-2}^{(j)}$ for a given value of j but are incident with the r remaining $(r-2)$-spaces $S_{r-2}^{(i)}$ $[i\neq j]$. By the method explained by C. A. Rupp * we can write down the equation of $W^{(j)}$ without difficulty. Write

* "The equation of V_{n-1}^{n-1} in S_n," *Bulletin of the American Mathematical Society*, Vol. 35 (1929), pp. 319-320.

$$\Phi_j \equiv \begin{vmatrix} -\Sigma a_i{}^{(j+1)}x_i & a^{(j+1)}_{j+2}x_{j+2} & \cdots & a^{(j+1)}_{j+r}x_{j+r} \\ a^{(j+2)}_{j+1}x_{j+1} & -\Sigma a_i{}^{(j+2)}x_i & \cdots & a^{(j+2)}_{j+r}x_{j+r} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a^{(j+r)}_{j+1}x_{j+1} & a^{(j+r)}_{j+2}x_{j+2} & \cdots & -\Sigma a_i{}^{(j+r)}x_i \end{vmatrix}$$

where $j+n \equiv j+n-r-1$ if $j+n>r$. Φ_j being divisible by x_j, we have

$$\Phi_j/x_j = 0$$

for the equation of $W^{(j)}$. We see at once that the equation of a general $V^{\,r}_{r-1}$ of $|\,V\,|$ is

$$\sum_{j=0}^{r} A_j\Phi_j \equiv A_0\Phi_0 + A_1\Phi_1 + \cdots + A_r\Phi_r = 0$$

where the A's are arbitrary constants. The $V^r{}_{r-1}$ whose equation is $\Phi_i = 0$ is degenerate, being composed of $W^{(j)}$ and the hyperplane $x_j = 0$.

3. *The birational transformation between S_r and R_r.* Letting $(y_0 : y_1 : \cdots : y_r)$ be the coördinates of a point in R_r, we have

$$y_0 : y_1 : \cdots : y_r = \Phi_0 : \Phi_1 : \cdots : \Phi_r$$

for the equations of the birational transformation of order r between S_r and R_r. If these equations are solved for the x's in terms of the y's, we obtain

$$x_0 : x_1 : \cdots : x_r = \Theta_0 : \Theta_1 : \cdots : \Theta_r$$

where Θ_j is the result of interchanging the superscripts and the subscripts of the a's and replacing x_i and x_{j+n} by y_i and y_{j+n} respectively in Φ_j.

In the r-space R_r are $r+1$ $(r-2)$-spaces whose equations are

$$R^{(j)}_{r-2}\colon \quad y_j = 0, \qquad \sum_{i=0}^{r} a_j{}^{(i)}y_i = 0.$$

The equation

$$\Theta_j/y_j = 0$$

represents the ruled V^{r-1}_{r-1} whose generators meet r of the $r+1$ $(r-2)$-spaces above but not the $R^{(j)}_{r-2}$ for a given value of j.

4. *The involutorial transformation.* We have hitherto imposed no restrictions upon the a's. If we put $a_i{}^{(j)} = a_j{}^{(i)}$, then we obtain

$$\Phi_0/x_0 = \Phi_1/x_1 = \cdots = \Phi_r/x_r, \qquad \Theta_0/y_0 = \Theta_1/y_1 = \cdots = \Theta_r/y_r.$$

This yields an identical transformation which is no interest.

Now suppose $a_i{}^{(j)} = -a_j{}^{(i)}$. This supposition causes the transformation

above to become involutorial. Letting $(x'_0 : x'_1 : \cdots : x'_r)$ be the coördinates of a point in the same r-space S_r, we have for the equations of this transformation

$$x'_0 : x'_1 : \cdots : x'_r = \Delta_0 : \Delta_1 : \cdots : \Delta_r$$

where Δ_j is the result of putting $a_j{}^{(i)} = -a_i{}^{(j)}$ in Φ_j or of putting $a_j{}^{(i)} = -a_i{}^{(j)}$ and replacing the y's by the x's in Θ_j.

This involutorial r-ic transformation is a very special case of the one effected by means of r given quadric hypersurfaces $Q^{(k)}$ $[k = 1, 2, \cdots, r]$. To a point P corresponds the point P', the intersection of the polar hyperplanes of P with respect to $Q^{(k)}$. The Jacobian locus, i. e., the locus of the points whose polar hyperplanes intersect in lines instead of points, is an $(r-2)$-dimensional variety, $J_{r-2}^{r(r-1)/2}$, of order $r(r-1)/2$ and the locus of lines so obtained is a ruled hypersurface, $V_{r-1}^{r^2-1}$, of order $r^2 - 1$.

In the present case, the $J_{r-2}^{r(r-1)/2}$ is composed of the $r + 1$ given $(r-2)$-spaces $S_{r-2}^{(j)}$ and the ruled variety $M_{r-2}^{(r+1)(r-2)/2}$ whose lines are incident with all the $r + 1$ given $(r-2)$-spaces and the $V_{r-1}^{r^2-1}$ is composed of the $r + 1$ $(r-1)$-ic hypersurfaces $W^{(j)}$. For the r quadric hypersurfaces $Q^{(k)}$ we may take any r of the $r + 1$ pairs of hyperplanes

$$x_j = 0, \qquad \sum_{i=0}^{r} a_i{}^{(j)} x_i = 0 \qquad\qquad [a_j{}^{(i)} = -a_i{}^{(j)}].$$

Or, we may consider the $r + 1$ degenerate quadric hypersurfaces $Q^{(j)}$

$$F^{(j)} \equiv x_j \sum_{i=0}^{r} a_i{}^{(j)} x_i = 0, \qquad [j = 1, 2, \cdots, r;\ a_j{}^{(i)} = -a_i{}^{(j)}].$$

The F's satisfy the identity

$$F^{(0)} + F^{(1)} + \cdots + F^{(r)} \equiv 0.$$

If $F_m{}^{(j)}$ is written for $\partial F^{(j)}/\partial x_m$, the Jacobian determinant of the F's is

$$D \equiv | F_m{}^{(j)} |,$$

which vanishes identically. Δ_j is the r-rowed determinant obtained from D by omitting the row and column containing $F_j{}^{(j)}$. As a matter of fact, any of the $r + 1$ r-rowed determinants formed from D with the column $F_j{}^{(0)}, F_j{}^{(1)}, \cdots, F_j{}^{(r)}$ removed may be taken as Δ_j.

THE UNIVERSITY OF CALIFORNIA.

INVOLUTORIAL SPACE CREMONA TRANSFORMATIONS DETERMINED BY NON-LINEAR NULL RECIPROCITIES.

By Edwin J. Purcell.

1. *Introduction.* This paper treats birational correspondences of space in which any two corresponding points are reciprocal with respect to a quadric polarity, or in a null-system. In these correspondences a general plane is transformed into a monoidal surface, where the term ***monoidal surface*** is used in the extended sense to mean any surface which has in common with a general line of a congruence of order one a single point which is non-singular for the congruence. Montesano * investigated these transformations synthetically for the non-involutorial cases.

The present paper arrives at many of his results analytically, and gives the equations of the transformations; it also gives the orders of the transformations, which Montesano did only for the case where a general plane is transformed into a monoidal surface, in the usual restricted meaning of this term. We are particularly concerned with the involutorial cases; these were not considered in Montesano's work.

These birational correspondences are generated in the following manner. Consider a plane Cremona transformation, C. T., of order n between the points of an arbitrary fixed plane y in space S and the points of an arbitrary fixed plane y' in a second space S'. Let each ray p of a congruence Q of order one in S be determined by its point of intersection with y, which is sent over by C. T. into a point of y', the latter point in turn determining a single ray p' of the congruence Q' in S'. This gives a birational correspondence X between the rays of Q and the rays of Q'. Any point P in S lies on a single ray p of Q which goes over by X into a single ray p' of Q'. A reciprocity Γ between the points of S and the planes of S' sends P over into a plane π' in S' which intersects p' in a point P', the correspondent of P in the birational correspondence $K = X.\Gamma$.

The congruences Q of order one which are considered are of three species: †

First species, when Q consists of a bundle of lines through a fixed point O of space;

* D. Montesano, "Sulle Reciprocita Birazionali Nulle dello Spazio," *Rendiconti della R. Accademia dei Lincei,* Vol. 4 (1888), pp. 583-590.

† D. Montesano, *loc. cit.*, p. 586.

Second species, when it consists of the lines of space which intersect a fixed line d and once a fixed curve Δ_μ of order μ having $\mu - 1$ points on d; (a limiting case of this arises from the planes of a pencil and the points D of its axis in $(\mu, 1)$ correspondence; any point P in space determines the plane of the pencil through it, and this determines D, thus P uniquely determines a line DP through it);

Third species, when it consists of the bisecants of a fixed twisted cubic curve Δ_3.

Montesano * said that when Q and Q' are both of the second species, to a ruled surface consisting of rays of Q intersecting an arbitrary fixed line of space S there corresponds, by X, a surface in Q' which he calls F'_n. This n means $n(\mu + 1)^2$, where the n is now the order of the plane Cremona transformation. The effect of common singular points of Q and of C. T. are not considered in his paper.

PART I. *The congruences Q and Q' both of the first species.*

2. Let the fixed plane y of space S be $x_4 = 0$, the fixed plane y' in S' be $x'_4 = 0$, the fixed vertex O of the bundle of rays Q be $(0, 0, 0, 1)$ in S, the fixed vertex O' of the bundle of rays Q' be $(0, 0, 0, 1)$ in S', the plane Cremona transformation C. T. between the points of $x_4 = 0$ and $x'_4 = 0$ be

$$\text{C. T.:}\quad \begin{aligned} \rho x'_1 &= \phi_n(x_1, x_2, x_3), \\ \rho x'_2 &= \psi_n(x_1, x_2, x_3), \\ \rho x'_3 &= \chi_n(x_1, x_2, x_3), \end{aligned} \qquad \text{C. T.:}\ h_1^{\alpha_1}, \cdots, h_r^{\alpha_r},$$

and the inverse of C. T. be

$$(\text{C. T.})^{-1};\ \tau x_1 = \phi'_n(x');\ (\text{C. T.})^{-1};\ h_1^{\beta_1}, \cdots, h_r^{\beta_r}, \text{ etc.}$$

Furthermore, let the correlation Γ between the points of S and the planes S' be

$$\Gamma:\ \sigma u'_i = x_i, \qquad (i = 1, 2, 3, 4.)$$

The equations of the birational transformation are

$$K:\quad \begin{aligned} &\lambda x'_1 = x_4\phi_n(x_1, x_2, x_3),\ \lambda x'_2 = x_4\psi_n(x_1, x_2, x_3), \\ &\lambda x'_3 = x_4\chi_n(x_1, x_2, x_3),\ \lambda x'_4 = -(x_1\phi_n + x_2\psi_n + x_3\chi_n). \end{aligned}$$

It is of order $(n + 1)$.

Similarly, the inverse transformation K^{-1} is

$$K^{-1}:\ \mu x_1 = x'_4\phi'_n(x'), \text{ etc.}$$

* D. Montesano, *loc. cit.*, p. 586.

Now suppose the two spaces S and S' coincide and that the tetrahedra of reference are identical. The number of isolated invariant points in the plane $x_4 = 0$ is $(n+2)$.* Therefore the number of invariant rays of the bundle O is $(n+2)$. In the coincident spaces Γ gives a quadric surface such that any point on it lies on its corresponding plane. Each invariant ray intersects the quadric surface in two points. Therefore there are $2(n+2)$ self-corresponding points for K when S and S' coincide.

PART II. *The congruences Q and Q' are both of the second species.*

3. *The spaces S and S' distinct.* Let the fixed line d be $x_1 = 0$, $x_2 = 0$ in space S and the parametric equations of the fixed curve Δ_μ in S be

$$\Delta_\mu: \quad \begin{array}{ll} x_1 = \phi(as+bt), & x_2 = \phi(cs+dt,) \\ x_3 = f_\mu(s,t), & x_4 = F_\mu(s,t). \end{array}$$

In the space S' take the fixed line d' to be $x'_1 = 0$, $x'_2 = 0$, and the fixed curve $\Delta_{\mu'}$ to have parametric equations of the same form as Δ_μ but with parameters s' and t'. Take the plane Cremona transformation to be the C. T. of section 2. Let the correlation Γ be the same as in section 2. Then

$$K: \quad \begin{aligned} x'_1 &= [\{f_\mu(U) - D\chi_n(W)\}\, x_3 + F_\mu(U)x_4]\, \phi_n(W), \\ x'_2 &= [\{f_\mu(U) - D\chi_n(W)\}\, x_3 + F_\mu(U)x_4]\, \psi_n(W), \\ x'_3 &= \{D\chi_n(W) - f_\mu(U)\}\, \phi_n(W)x_1 + \{D\chi_n(W) - f_\mu(U)\}\, \psi_n(W)x_2 \\ &\qquad + F_\mu(U)\chi_n(W)x_4, \\ x'_4 &= \{\phi_n(W)x_1 + \psi_n(W)x_2 + \chi_n(W)x_3\}\, F_\mu(U), \end{aligned}$$

where

$$\begin{aligned} (U) &\equiv [\, b\psi_n(W) - d\phi_n(W),\ c\phi_n(W) - a\psi_n(W)], \\ (W) &\equiv [\{F_\mu(N) - Cy_4\}\, y_1,\ \{F_\mu(N) - Cy_4\}\, y_2,\ \{y_3F_\mu(N) - y_4f_\mu(N)\}], \\ (N) &\equiv (by_2 - dy_1,\ cy_1 - ay_2), \\ C &\equiv (bc - ad) \prod_{i=1}^{\mu-1} \{t_i(by_2 - dy_1) - s_i(cy_1 - ay_2)\}, \end{aligned}$$

and s_i, t_i $(i = 1, 2, \cdots, \mu - 1)$ are the values of the parameters s, t of Δ_μ at the $\mu - 1$ points where Δ_μ intersects the fixed line d.

The transformation K is of order $n(\mu+1)^2 + 1$.

Similarly, the inverse transformation K^{-1} is found.

4. *Involution.* If Q and Q' coincide in the same space $S \equiv S'$ and Γ, as before, is a quadric polarity, then the birational transformation K will be

* H. Hudson, *Cremona Transformations* (1927), p. 78.

an involution if C. T. is a plane Cremona involution. Here the direct and inverse transformations C. T., being involutorial, have the same form.

The involution K is the same as the transformation K defined in section 3 except that the ϕ_n, ψ_n, χ_n are the functions that appear in the plane Cremona involution C. T.

The order of the involution K is $n(\mu+1)^2+1$.

The inverse of this involution is the same as K but with x_i and x'_i interchanged $(i=1, 2, 3, 4)$.

In this paper, Γ is considered to be the quadric polarity whose equations appear in section 2. The problems which occur when Γ is a null-system will be mentioned as they arise.

5. *Fundamental points of K where the plane of the Cremona involution C. T. cuts Δ_μ and d.* The directrix d of Q intersects $x_4=0$ in the point $(0,0,1,0)$ which is a fundamental point for K. Now any generator of the projecting cone Σ with vertex $(0,0,1,0)$ and base curve Δ_μ will be a ray of Q intersecting $x_4=0$ in $(0, 0, 1, 0)$. This cone is of order μ. If $(0, 0, 1, 0)$ is a regular point for the plane Cremona involution C. T., then it goes over, by C. T., into another point δ of $x_4=0$, also regular for C. T. If the point δ does not lie on Δ_μ, it determines one ray ρ of Q. The projecting cone Σ of order μ goes over, by K, into ρ. Any point on Σ goes over into a point on ρ. Since C. T. is an involution, δ goes over into $(0, 0, 1, 0)$ and therefore K transforms ρ into the cone Σ. Any point on ρ goes over into the plane curve of order μ which is the intersection of the polar plane of the point and the cone Σ.

The curve Δ_μ intersects $x_4=0$ in μ points, $m_1, m_2, \cdots, m_\mu$. The whole pencil of lines through m_i intersecting d are rays of Q. If m_i is a regular point for C. T., it goes over by C. T. into another point σ_i, regular for C. T. If σ_i is neither on d nor Δ_μ, it determines a single ray ρ_i of Q. Then any point on the plane determined by m_i and d will go over by K into a point on ρ_i. Since C. T. is an involution, ρ_i goes over into the whole pencil of lines with vertex m_i lying in the plane of m_i and d. A point of ρ_i is transformed by K into a line of the plane of m_i and d.

Should the point δ lie on Δ_μ, say at m_k, then the cone Σ goes over into the plane of m_k and d, and vice versa. Any point on Σ will go over into a line on the plane of m_k and d, and any point on this plane will go over into a plane curve of order μ on Σ.

Should any of the points m_i, say m_j, go over by C. T. into another m_i, say m_k, then the plane of m_j and d will go over into the plane of m_k and d

and vice versa. Any point on one of these two planes will go over into a line on the other.

Consider a line M_i joining m_i to $(0, 0, 1, 0)$. By C. T., M_i goes over into a curve of order n in $x_4 = 0$, which determines a ruled surface of order $n(\mu + 1)$ of rays of Q. Any point on M_i goes over into a plane curve of order $n(\mu + 1)$, the intersection of the polar plane of the point with this ruled surface. Any point on the ruled surface goes over into a point on M_i.

6. *The invariant locus for the involution K.* Suppose there is a curve I, of order i, of invariant points for the plane Cremona involution C. T. in $x_4 = 0$. The ruled surface R_I of rays of Q which intersect I will be invariant as a whole and is of order $(\mu + 1)i$. Any generator of this ruled surface intersects the quadric surface q, which is the locus of self-conjugate elements in the polarity Γ. It follows that every point of the curve of intersection of the ruled surface R_I and the quadric q is invariant in the involution K. This invariant curve for K is of order $2(\mu + 1)i$, if I does not intersect d or Δ_μ. If Γ is a null polarity, then every point of each line of Q through I is invariant; hence the invariant surface consists of the ruled surface of Q on I. If I does intersect d or Δ_μ, the order of K is reduced, as will be seen in the next section.

In addition to I, there may be isolated invariant points for C. T. in $x_4 = 0$. Each of these determines a ray of Q which is invariant as a whole for K, and intersects the quadric q, invariant for Γ, in two points. These two points are isolated invariant points of the involution K.

7. *Reduction of the order of the involution K.*

THEOREM 1. *If the intersection of the directrix d with the plane of the Cremona involution C. T. is invariant for C. T., then the projecting cone of Δ_μ from this point, to multiplicity μ is a factor of the equations of the involution K, whose order is thereby reduced by μ^2.*

THEOREM 2. *Any point of intersection of the curve Δ_μ with the plane of the Cremona involution C. T., which is invariant for C. T., causes the plane determined by this point and the directrix d to factor out of the equations of the involution K, whose order is reduced by one for each such point.*

In addition to the fundamental points of K in $x_4 = 0$ arising from the intersection of d or Δ_μ with $x_4 = 0$, there are fundamental points of the C. T. Consider an F-point O_a of C. T., of multiplicity α. If it does not lie on d or Δ_μ, it determines a single ray p of Q. But O_a goes over, by C. T., into a curve j_a of order α in $x_4 = 0$. The rays of Q which intersect j_a form a ruled

surface R of order $(\mu + 1)\alpha$. In K, the ray p through O_α goes over into the ruled surface R, and any point on p goes over into a plane curve, of order $(\mu + 1)\alpha$, which is the intersection of R and of the polar plane of the point. Any ray of Q on R goes over into the whole ray p, and any point on R goes over into a point on p.

THEOREM 3. *If the intersection of the directrix d with the plane of the Cremona involution C. T. is an α-fold fundamental point for C. T., then the projecting cone of the curve Δ_μ from this point, to multiplicity $(\mu + 1)\alpha$, which is the order of the ruled surface of rays of Q intersecting the principal curve for C. T. corresponding to the fundamental point considered, factors from the equations of the involution K, whose order is thereby reduced by $\mu(\mu + 1)\alpha$.*

THEOREM 4. *Any intersection of Δ_μ with the plane of the Cremona involution C. T. which is a β-fold fundamental point for C. T. causes the plane of this point and of the directrix d to factor out from the equations of the involution K to multiplicity $(\mu + 1)\beta$, which is the order of the ruled surface of rays of Q which intersect the principal curve of C. T. corresponding to the β-fold fundamental point considered. This reduces the order of K by $(\mu + 1)\beta$.*

9. *A limiting case of the involution K.* A limiting case of the congruence Q of the second species arises from the planes of a pencil and the points D of its axis in $(\mu, 1)$ correspondence. Any point P in space determines the plane of the pencil through it, and this determines D. In this manner P uniquely determines a line DP through it. Let the axis d of the pencil of planes be $x_1 = 0$, $x_2 = 0$. Take as the relation giving the $(\mu, 1)$ correspondence between the planes of the pencil and the points $(0, 0, z_3, z_4)$ of the axis d

$$\rho z_3 = f_\mu(y_1, y_2), \quad \rho z_4 = F_\mu(y_1, y_2).$$

Using the C. T. in $x_4 = 0$ and Γ of section 2 in the manner of section 4, the equations of the involution K are found to be

$$K: \quad \begin{aligned} x'_1 &= [x_3 f_\mu \{\phi_n(W), \psi_n(W)\} + x_4 F_\mu \{\phi_n(W), \psi_n(W)\}]\, \phi_n(W), \\ x'_2 &= [x_3 f_\mu \{\phi_n(W), \psi_n(W)\} + x_4 F_\mu \{\phi_n(W), \psi_n(W)\}]\, \psi_n(W), \\ x'_3 &= [x_3 f_\mu \{\phi_n(W), \psi_n(W)\} + x_4 F_\mu \{\phi_n(W), \psi_n(W)\}]\, \chi_n(W), \\ &\quad - \{x_1\phi_n(W) + x_2\psi_n(W) + x_3\chi_n(W)\}\, f_\mu \{\phi_n(W), \chi_n(W)\} \\ x'_4 &= - \{x_1\phi_n(W) + x_2\psi_n(W) + x_3\chi_n(W)\}\, F_\mu\{\phi_n(W), \psi_n(W)\}, \end{aligned}$$

where $(W) \equiv [x_1 F_\mu(x_1, x_2),\ x_2 F_\mu(x_1, x_2),\ x_3 F_\mu(x_1, x_2) - x_4 f_\mu(x_1, x_2)]$. K

is of order $n(\mu+1)^2+1$, and contains d to multiplicity $\mu(\mu+1)n$. Although this involution is a limiting case of the one in section 4, it cannot be obtained by specialization of the coefficients of the latter. The only singular point of this congruence is $(0, 0, 1, 0)$ where d meets $x_4=0$. In $x_4=0$ are μ lines, all passing through this point. Let the ray p determined by (y) meet $x_4=0$ in an r-fold F-point of C. T. Its image in C. T. is a rational curve C_r of order r in $x_4=0$. The ruled surface on C_r belonging to the congruence contains the μ lines each r-fold, and therefore its order is $r(\mu+1)$. No generator can meet any other one except on d. From any point of d, C_r determines k_r. In the congruence all the lines must lie in μ planes. Therefore d is r μ-fold. A point on the given line has for image a plane section of this surface, hence the line is of multiplicity $(\mu+1)r$.

PART III. *The congruences Q and Q' are both of the third species.*

10. *The spaces S and S' distinct.* Let the parametric equations of the fixed cubic curve Δ_3 in the first space S be

$$\Delta_3:\quad \begin{aligned} x_1&=\lambda^2(\lambda-\mu),\ x_2=\lambda^2\mu,\\ x_3&=\mu^2(\lambda-\mu),\ x_4=\lambda\mu(\lambda-\mu),\end{aligned}$$

and let the plane of the Cremona transformation C. T. be $x_4=0$. Similarly, in the second space S' let the congruence Q' be determined by the fixed cubic curve

$$\Delta'_3:\quad \begin{aligned} x'_1&=\lambda'^2(\lambda'-\mu'),\ x'_2=\lambda'^2\mu',\\ x'_3&=\mu'^2(\lambda'-\mu'),\ x'_4=\lambda'\mu'(\lambda'-\mu').\end{aligned}$$

Let the plane Cremona transformation C. T. be that of Part 1, § 2. The correlation Γ is $u'_i=x_i$, $(i=1, 2, 3, 4)$. Putting x_i for y_i $(i=1, 2, 3, 4)$, the equations of K are

$$K:\quad \begin{aligned} x'_1&=\phi_n(S)\ [\phi_n(S)\psi^2_n(S)x-2\chi_n(S)\{2\phi_n(S)\chi_n(S)\\ &\qquad +\psi_n(S)\chi_n(S)-\phi_n(S)\psi_n(S)\}x_3+2\phi_n(S)\psi_n(S)\chi_n(S)x_4],\\ x'_2&=\psi_n(S)\ [-\phi^2_n(S)\psi_n(S)x_1-\chi^2_n(S)\ \{2\phi_n(S)+\psi_n(S)\}\ x_3\\ &\qquad +\phi_n(S)\psi_n(S)\chi_n(S)x_4],\\ x'_3&=\chi_n(S)\ [2\phi_n(S)\ \{2\phi_n(S)\chi_n(S)+\psi_n(S)\chi_n(S)-\phi_n(S)\psi_n(S)\}\ x_1\\ &\qquad +\psi_n(S)\chi_n(S)\ \{2\phi_n(S)+\psi_n(S)\}\ x_2+2\phi_n(S)\psi_n(S)\chi_n(S)x_4],\\ x'_4&=-\phi_n(S)\psi_n(S)\chi_n(S)\ [2\phi_n(S)x_1+\psi_n(S)x_2+2\chi_n(S)x_3],\end{aligned}$$

where $(S)\equiv(BC, -AC, AB)$,

$$\begin{aligned} A&\equiv y_2y_4-y_2y_3-y_3y_4;\ B\equiv 2y_1y_3-y_1y_4-y_3y_4,\\ C&\equiv y_1y_4+y_2y_4-y_1y_2;\ D\equiv 2y_1y_3+y_2y_3-y_1y_2.\end{aligned}$$

K is of order $16n+1$.

Similarly the inverse transformation K^{-1} can be found.

11. *Involution.* Let Δ_3 and Δ'_3 coincide in the same space $S' \equiv S$. Then K is an involution if C. T. is a plane Cremona involution in $x_4=0$. Let C. T. be an involution, $x'_1 = \phi_n(x)$, etc. The involution K has equations of the form of K in section 10, where A, B, C, D, and S are as defined in section 10, and ϕ_n, ψ_n, χ_n are as in the plane Cremona involution C. T.

The order of K is $16n+1$.

12. *Fundamental points of the involution K where the plane of the Cremona involution C. T. cuts Δ_3.* The fixed curve Δ_3 of Q intersects $x_4=0$ in the three points $O_1 \equiv (1, 0, 0, 0)$, $O_2 \equiv (0, 1, 0, 0)$, and $O_3 \equiv (0, 0, 1, 0)$, which are fundamental points for K. Any ray of the quadric cone q_1 having one of these points, say O_1, as vertex and the curve Δ_3 as base curve is a ray of Q intersecting $x_4=0$ in O_1. If O_1 is a regular point for C. T., it goes over by C. T. into another point J_1 of $x_4=0$, also regular for C. T. If the point J_1 is not on Δ_3, it determines a single ray σ of Q. The quadric cone q_1 goes over, by K, into the single ray σ. Any point on q_1 goes over into a point on σ. Since C. T. is an involution, J_1 goes over into O_1 and therefore K transforms σ into the cone q_1. Any point on σ goes over into a conic on q_1 which is the intersection of the polar plane of the point, by Γ, and of the cone q_1.

Should the point J_1, image of O_1 by C. T., be either of the remaining intersections of Δ_3 with $x_4=0$, say O_2, then the cone q_1 goes over, by K, into the quadric cone q_2 with vertex O_2 and base curve Δ_3, and vice versa. In fact, any generator of one of these cones goes over into the whole of the other cone.

Consider the line M joining two of the intersections of Δ_3 with $x_4=0$. By C. T. it goes over into a curve of order n in $x_4=0$, which determines a ruled surface R, of order $4n$, of rays of Q which intersect it. This particular generator M of cones q_1 and q_2 goes over into the whole ruled surface R, of order $4n$, by K. Any point on M goes over into the plane curve of order $4n$, intersection of R, with the polar plane of the point in Γ. R goes over, by K, into the two cones with vertices on the intersections of M with Δ_3, and with Δ_3 as base curve. Any point on R, other than the two image points of O_1 and O_2, goes over into a point on M.

13. *The invariant locus for the involution K.* If there is a curve I, of order i, of invariant points for C. T. in $x_4=0$, then the ruled surface R_I of rays of Q which intersect I will be invariant as a whole for K and is of order

$4i$. Any generator of this ruled surface intersects the quadric surface, locus of self-conjugate elements in Γ. Therefore every point of the curve of intersection of this quadric and the ruled surface R_I is invariant in the involution K. This invariant curve is of order $8i$, if I does not intersect Δ_3. If Γ is a null-system, then every point of each line of Q through I is invariant, hence the invariant locus consists of the ruled surface of Q on I. If I does intersect Δ_3, the order of K is reduced as is shown in the next section.

Now there may also be isolated invariant points for C. T. in $x_4 = 0$. Each of these determines a ray of Q which is invariant as a whole for K. It intersects the self-polar quadric in two points which will be isolated invariant points for the involution K.

14. *Reduction of the order of the involution K.*

THEOREM 1. *Any intersection of Δ_3 with the plane of C. T. which is an invariant point for C. T. causes the projecting cone of Δ_3 from that point, to multiplicity two, to factor out of the equations of K, whose order is thus reduced by four.*

THEOREM 2. *Any intersection of Δ_3 with the plane of C. T. which is an α-fold fundamental point for C. T. causes the projecting cone of Δ_3 from that point, to multiplicity 4α, to factor out of the equations of K, whose order is thus reduced by 8α.*

COROLLARY. *The order of the involution K can always be reduced so as to be not greater than $8n - 7$.*

Proof. Let Δ_3 intersect the plane of C. T. in the three F-points of highest multiplicity for C. T. From Noether's inequality, the sum of the multiplicities of the three highest F-points of C. T. exceeds the degree n, provided $n > 1$. Then the order $16n + 1$ of K is reduced by at least $8(n + 1)$.

While a proper placing of the F-points of the Jonquieres C. T. gives the greatest reduction of the order of K when the congruence Q is of the second species, here quite the reverse is true. For, only in the case of the Jonquieres or Symmetric C. T.'s can the sum of the multiplicities of the three highest F-points be as little as $n + 1$, and consequently the order of K is reduced by more than $8(n + 1)$.

CORNELL UNIVERSITY.

AN APPLICATION OF THE DEDEKIND CUT NOTION TO INTEGRATION.*

By E. R. Hedrick and W. M. Whyburn.

In this paper, the authors give a simple method of defining the integral of a bounded function by means of a device similar to the well known Dedekind Cut in the theory of real numbers. The basal functions used are the so-called *simple functions* (or step-functions) whose integrals are but the sums of the areas of finite sets of rectangles. A first procedure, outlined below in connection with class $C(M)$ leads to a definition of integration which includes all bounded Riemann-integrable functions. An extension described in the treatment of class $D(M)$ makes possible the definition of the integral of any bounded Lebesgue-integrable function. This is accomplished without the aid of the usual theorems on the measure of a point set and it may be employed, once the Lebesgue integral is defined, to entirely replace that theory. In conclusion the possibility of extension to unbounded functions is noted. The work of the paper is confined to the real domain. All sets and collections used are understood to contain at least one element.

2. *Definitions.* Let X: $a \leqq x \leqq b$ be a finite portion of the real axis and let M be a set of points belonging to X. M is said to be a δ-*set*,† where δ is a positive real number, if there exists an at most countably infinite collection of sub-intervals of X that covers ‡ M and is such that the sum of the lengths of the intervals of this collection does not exceed δ. Clearly a δ-*set* is also an ϵ-*set* if ϵ is greater than δ. If a set M is a δ-set for each δ, $\delta > 0$, the set M is said to be a *null set*. A property that holds for all points of a set M with the possible exception of points of a sub-set of M which is a null set, is said to hold on M_0.

Collection of functions everywhere dense on M. A collection $[f(x)]$ of real functions on M is said to be *everywhere dense* on M if for each point $x = p$ of M, the set of numbers $[f(p)]$ is everywhere dense on the real number axis.

---

* Read before the American Mathematical Society, August 31, 1932.

† This, of course, is a portion of the notion of *outer measure*.

‡ A set of intervals on X is said to cover M if each point of M that is interior to X is interior to some one of the intervals of the collection and if M contains $x = a$ or $x = b$, then this point is an end point of some interval of the collection.

*Simple function.** A function $\phi(x)$ is said to be a *simple function* on X if there are $n+1$ points: $a = x_0 < x_1 < \cdots < x_n = b$ such that $\phi(x) = \phi_i$, a constant, on X_i: $x_i < x < x_{i+1}$. When it is desired that $\phi(x)$ be defined at all points of X, we understand $\phi(x_i) = \phi_i$, $(i = 0, 1, \cdots, n-1)$, $\phi(b) = \phi_{n-1}$. Let K denote the collection of all simple functions on X.

The Dedekind number cut.† Let G be a set of real numbers that is everywhere dense ‡ on the real number axis. A separation of G into two sets, $G = G_1 + G_2$, in such a way that $g_1 \leqq g_2$, where g_1 is any number in G_1 and g_2 is any number in G_2, is said to constitute a *Dedekind number cut.* Such a cut determines a unique number g which has the property $g_1 \leqq g \leqq g_2$ for each g_1 in G_1 and each g_2 in G_2. The number g is said to be *defined by* the cut and may not belong to the set G.

Cut in K on M. Let M be a subset of X and let G be a countable sub-collection of K that is everywhere dense on M_0. A separation of G into two sub-collections, $G = G_1 + G_2$, in such a way that $g_1(x) \leqq g_2(x)$ on M_0, where $g_1(x)$ is any simple function in G_1 and $g_2(x)$ is any simple function in G_2, is said to constitute a *cut in K on M.* Such a cut determines a unique function $f(x)$ on M_0 which has the property $g_1(x) \leqq f(x) \leqq g_2(x)$ on M_0 for each $g_1(x)$ in G_1 and each $g_2(x)$ in G_2. The function $f(x)$ so determined is said to be defined on M_0 by a cut in K on M.

The class $C(M)$. The collection of all functions $f(x)$ which are defined on M_0 by cuts in K on M and which are bounded on M_0 is called the class $C(M)$, where M is a sub-set of X.

Enlargement of a collection of simple functions. A countable collection $[\phi_i(x)]$ of simple functions is said to be *enlarged above* [*enlarged below*] if it is enlarged by the inclusion of all simple functions obtained in the following manner:

Let $X_{i1}, X_{i2}, \cdots, X_{in_i}$, be the sub-intervals on which $\phi_i(x)$ has constant values. Let $h_{ijr}(x) = \phi_i(x)$ on $X - X_{ij}$, $h_{ijr}(x) = r$ on X_{ij}, where r is a rational number greater than [less than] the value of $\phi_i(x)$ on X_{ij}. Let $[h_{ijr}(x)]$ be the collection of simple functions obtained when r ranges over all rational numbers that exceed [are exceeded by] the value of $\phi_i(x)$ on X_{ij}, $(j = 1, 2, \cdots, n_i)$, $(i = 1, 2, \cdots)$. The enlarged collection is $[\phi_i(x)] + [h_{ijr}(x)]$ and is made up of a countable set of simple functions.

* See F. Riesz, *Acta Mathematica*, Vol. 42 (1920), pp. 191-205.

† We state a form of this cut that is useful in our paper. The set G may be countable.

‡ It is not essential that G be everywhere dense on the real axis for a particular cut (simply dense on some interval with g on its interior). Similar statements apply to our function cuts.

The class $D(M)$. A function $f(x)$, defined and bounded on M, is said to belong to *class* $D(M)$ if it has the property that for each $\delta > 0$, there is a subset $M(\delta)$ of M such that $M - M(\delta)$ is a δ-set and such that the function $f(x, \delta)$ that equals $f(x)$ on $M(\delta)$ and equals zero at points of $X - M(\delta)$ is of class $C(X)$.

3. *Theorems on functions of $C(M)$ and $D(M)$.*

THEOREM I. *For each function $f(x)$ in $C(M)$ there exist uniformly bounded sequences $[h_i(x)]$ and $[g_i(x)]$ of simple functions such that $\lim_{i\to\infty} h_i(x) = f(x)$, $\lim_{i\to\infty} g_i(x) = f(x)$, on M_0 and, furthermore, $h_i(x) \geqq f(x) \geqq g_i(x)$ for each x on M_0, $(i = 1, 2, \cdots)$.*

Proof. Since $f(x)$ is in $C(M)$, it is defined on M_0 by a cut in K on M. Let G_1 and G_2 be the collections of simple functions used in this cut. Let E denote the null set composed of the points of M at which $G_1 + G_2$ fails to define $f(x)$ together with all points of M that are points of discontinuity for the simple functions of $G_1 + G_2$. Let $H = M - E$. Let k be a positive integer and let X be divided into k intervals of equal length. On each sub-division let $h_k(x)$ be the upper bound and $g_k(x)$ be the lower bound of $f(x)$ on the subset of H that belongs to that sub-division. $[h_k(x)]$ and $[g_k(x)]$, the collections of simple functions * obtained when $k = 1, 2, 3, \cdots$, are the sequences desired for the theorem. These sequences are uniformly bounded since $f(x)$ is bounded on M. The inequality $g_i(x) \leqq f(x) \leqq h_i(x)$, $(i = 1, 2, 3, \cdots)$, on H follows immediately from the manner of construction of $h_i(x)$ and $g_i(x)$. Let $x = p$ be any point of H, and let $\epsilon > 0$ be an arbitrarily assigned number. Let $q_1(x)$ and $q_2(x)$ be picked from G_1 and G_2, respectively, so that $0 \leqq q_2(p) - q_1(p) < \epsilon$. Since $x = p$ belongs to H, $q_1(x)$ and $q_2(x)$ are continuous at $x = p$ and we can pick an interval J with $x = p$ on its interior such that $q_1(x)$ and $q_2(x)$ are constant on J. The upper and lower bounds of $f(x)$ on the subset of H that belongs to J differ by less than ϵ. We may choose an index n such that for all $k \geqq n$, the sub-division of X that contains $x = p$ will be a sub-interval of J and hence $0 \leqq h_k(p) - g_k(p) < \epsilon$. Hence $\lim_{k\to\infty} h_k(p) = \lim_{k\to\infty} g_k(p) = f(p)$. This completes the proof of Theorem I.

THEOREM II. *A necessary and sufficient condition that a bounded function $f(x)$ on M belong to $C(M)$ is that $f(x)$ be continuous on a set of points H, where $M - H$ is a null set.*

* $h_k(x)$ and $g_k(x)$ are defined to be zero on sub-divisions that contain no points of H.

Proof. (*Necessity*): Let $f(x)$ be in $C(M)$ and let H, $[h_i(x)]$, $[g_i(x)]$, be defined as they were in the proof of Theorem I. It follows directly from $\lim_{k\to\infty} h_k(x) = \lim_{k\to\infty} g_k(x) = f(x)$ on H that $f(x)$ is continuous on H.

Proof. (*Sufficiency*): Let $f(x)$ be bounded on M and continuous on H, where $M - H$ is a null set. Let k be a positive integer and let X be subdivided into k subdivisions of equal length. On each of these subdivisions, let $g_k(x)$ and $h_k(x)$ be the lower and upper bounds, respectively, of $f(x)$ on the subset of H that belongs to the subdivision. If the subdivision contains no points of H, let $g_k(x) = h_k(x) = 0$ on that subdivision. Let $[g_k(x)]$ and $[h_k(x)]$ be the sequences obtained when $k = 1, 2, 3, \cdots$. Let G_1 denote the collection of simple functions obtained when $[g_k(x)]$ is *enlarged below* and let G_2 be the collection obtained when $[h_k(x)]$ is *enlarged above.* The cut $G_1 + G_2$ in K on M defines the function $f(x)$ on H.

THEOREM III. *The class $C(M)$ is closed in the sense that any bounded function $F(x)$ that is defined on M_0 by a cut* in $C(M)$ on M belongs to $C(M)$.*

Proof. Let† $C = C_1 + C_2$ be a cut in $C(M)$ that defines $F(x)$ on M_0, where the elements of C are functions of $C(M)$ and a cut in $C(M)$ is defined in a manner entirely analogous to that used in defining a *cut in K on M.* Let $f_1(x)$ be a function of C_1 and $f_2(x)$ a function of C_2. The functions $f_1(x)$ and $f_2(x)$ are in $C(M)$ and are defined on M_0 by cuts $G_1 + G_2$ and $H_1 + H_2$, respectively, in K on M. Let W be the subset of M on which all of the functions of C and $F(x)$ are defined by their cuts. The set $M - W$ is a null set since the collection C is countable. Let $[G_1]$, $[G_2]$, $[H_1]$, $[H_2]$, be collections of simple functions composed of all of the functions of G_1, G_2, H_1, H_2, respectively, for all functions $f_1(x)$ and $f_2(x)$ of C_1 and C_2, respectively. The separation $[G_1] + [H_2]$ is a cut in K on M that defines $F(x)$ on W. This follows immediately when account is taken of the construction of W and the method of selection used for $[G_1]$ and $[G_2]$.

THEOREM IV. *The class $C(M)$ is a sub-class of $C(M')$, where M' is a subset of M.*

Proof. Any function $f(x)$ of $C(M)$ is defined on M_0 by a cut in K on M. This same cut defines $f(x)$ on M'_0 and hence $f(x)$ is in $C(M')$.

---

* The definition of such a cut is entirely analogous to that of a *cut in K on M.*

† The notation is chosen so that the elements of C_1, G_1, H_1, do not exceed those of C_2, G_2, H_2, respectively.

4. *Integration.*

Integral of simple function. Let $\phi(x)$ be a simple function on X with sub-division points $a = x_0 < x_1 < \cdots < x_n = b$ and let $\phi(x) = \phi_i$ on $x_i < x < x_{i+1}$. The *integral of* $\phi(x)$ or *the area under* $\phi(x)$ on X is defined to be $A[\phi] = \int_a^b \phi(x)\,dx = \sum_{i=0}^{n-1} \phi_i\,[x_{i+1} - x_i]$.

Integral of a function in class $C(X)$. Let $G_1 + G_2$ be any cut in K that defines $f(x)$ on X_0 and has the further property that the number sets A_1 and A_2, made up of the integrals of the simple functions in G_1 and G_2, respectively, define a Dedekind number cut. The integral, $\int_a^b f(x)\,dx$, of $f(x)$ on X is defined to be the number A determined by the Dedekind number cut $A_1 + A_2$.

THEOREM V. *If $f(x)$ is any function in $C(X)$, the integral $\int_a^b f(x)\,dx$ exists and is unique.*

Proof. We first show the existence of the integral. Let $[g_i(x)]$ and $[h_i(x)]$ be the uniformly bounded sequences of simple functions whose existence was established in Theorem I. Let N be a bound for these sequences and let G and H, respectively, denote the collections obtained when $[g_i(x)]$ is *enlarged below* and $[h_i(x)]$ is *enlarged above.* $G + H$ is a cut in K that defines $f(x)$ on X_0. Let $A_1 = [a_1]$ and $A_2 = [a_2]$, respectively, be the collections of area numbers for the functions of G and H. It follows from the definitions of collections G and H that each number a_1 in A_1 is less than or equal to each number a_2 in A_2 and that the numbers of $A_1 + A_2$ are everywhere dense on any number interval if both of the ends of this interval belong to A_1 or both belong to A_2. Let $\epsilon > 0$ be arbitrarily assigned and for each index i, let E_i be the finite set of sub-intervals of X on which $h_i(x) - g_i(x) > \epsilon/[4(b-a)]$. The set of points E common to infinitely many of the sets E_i is a null set since $\lim_{i\to\infty} [h_i(x) - g_i(x)] = 0$ on X_0. A theorem of F. Riesz * shows that $\lim_{i\to\infty} L_i = 0$, where L_i is the sum of the lengths of the intervals of E_i. Choose an index i so that $L_i < \epsilon/(4N)$. Then

$$(1) \qquad A[h_i] - A[g_i] < \epsilon(b-a)/[4(b-a)] + \epsilon(2N)/(4N) < \epsilon.$$

Since $A[h_i]$ belongs to A_2 and $A[g_i]$ belongs to A_1, we have now shown that $A_1 + A_2$ is everywhere dense on the real number axis and defines a Dedekind number cut. Let A be the number defined by this cut, then $A = \int_a^b f(x)\,dx$.

---

* *Loc. cit.*, page 195.

We now show that the integral of $f(x)$ on X is unique. Let $J_1 + J_2$ be any cut in K on X that defines $f(x)$ on X_0 and let $j_1(x)$ and $j_2(x)$, $j_1(x) \leqq j_2(x)$, be any functions of J_1 and J_2, respectively. The inequalities $j_1(x) \leqq h_i(x)$, $j_2(x) \geqq g_i(x)$ hold on X for $i = 1, 2, \cdots$ since a violation of one of these at a point of X would mean its violation on an interval of positive length which, in turn, would violate $g_i(x) \leqq f(x) \leqq h_i(x)$, $j_1(x) \leqq j_2(x)$ on X_0. Hence $\int_a^b j_1(x)\,dx \leqq A \leqq \int_a^b j_2(x)\,dx$ for each function $j_1(x)$ in J_1 and each function $j_2(x)$ in J_2. Hence if the cut $J_1 + J_2$ yields an $\int_a^b f(x)\,dx$, this integral must be equal to A. This completes the proof of Theorem V.

Corollary. *If $f(x)$ is in $C(X)$ and $[g_i(x)]$, $[h_i(x)]$ are the sequences of simple functions whose existence was established in Theorem I, then*

$$\lim_{i\to\infty} \int_a^b g_i(x)\,dx = \lim_{i\to\infty} \int_a^b h_i(x)\,dx = \int_a^b f(x)\,dx.$$

Definition of integral of a function in $D(M)$. Let $f(x)$ belong to $D(M)$, $|f(x)| < N$ on M, and $A(\delta) = \int_a^b f(x, \delta)\,dx$. Let A'_1 and A'_2 be the collections of numbers $A(\delta) - N\delta$ and $A(\delta) + N\delta$, respectively, obtained when δ takes all values between zero and $b - a$. Let A_1 denote A'_1 together with all real numbers less than the lower bound of the numbers in A'_1 while A_2 denotes A'_2 together with all real numbers greater than the upper bound of the numbers in A'_2. The separation $A_1 + A_2$ defines a Dedekind number cut and determines a unique * number A that is defined to be the integral, $\int_{(M)} f(x)\,dx$, of $f(x)$ on M.

Theorem VI. *The class $C(X)$ contains all bounded Riemann-integrable functions on X and is contained in the class of all bounded Lebesgue-integrable functions on X.*

Proof. Let $f(x)$ be bounded and Riemann integrable on X. The set of points of X at which $f(x)$ is discontinuous form a null † set, E. The function $f(x)$ is therefore continuous on $X - E$ and hence, by Theorem II, belongs to $C(X)$.

---

* The uniqueness of the integral follows when account is taken of the fact that two different choices of $f(x, \delta)$, for a given δ, would cause a variation in $A(\delta)$ that does not exceed $2N\delta$. This difference can be made arbitrarily small by choosing δ sufficiently small.

† See Lebesgue, *Annali di Matematico* (3), Vol. 7, page 254.

Let $f(x)$ belong to $C(X)$. Theorem I shows that $f(x)$ is the limit function on X_0 of a sequence of simple functions and a theorem of W. M. Whyburn's * shows that $f(x)$ is Lebesgue integrable on X.

We now give two examples. The first of these shows that $C(X)$ contains functions that are not Riemann integrable while the second shows that $C(X)$ does not contain all bounded Lebesgue integrable functions on X.

Example I. Let $f(x) = 0$ for rational values of x on X: $0 \leqq x \leqq 1$ and $f(x) = 1$ at all other points of X. The function $f(x)$ is not Riemann integrable on X since it is discontinuous at each point of this interval. It is in class $C(X)$ since it is continuous on X_0, where X_0 may be taken as the set of irrational points on X.

Example II. Let the rational points on X: $0 \leqq x \leqq 1$ be covered by a countable set of intervals of total length less than $1/2$ and let G be the set of all points that are interior to intervals of this set. Let $f(x) = 0$ at points of G, $f(x) = 1$ at points of $X - G$. The function $f(x)$ is bounded and measurable and is therefore Lebesgue integrable on X. It is not in $C(X)$, however, since there is no null set E such that $f(x)$ is continuous on $X - E$. This follows since any set $X - E$, where E is a null set, must contain subsets of G and $X - G$ both of which are everywhere dense on X.

THEOREM VII. *If M is a measurable point set, the class $D(M)$ is identical with the collection of all bounded and measurable functions on M.*

Proof. Let $f(x)$ be defined to be zero at all points of $X - M$. If $f(x)$ is bounded and measurable on M, it is bounded and measurable on X and a theorem of W .M. Whyburn's † states that $f(x)$ is the limit function on X_0 of a sequence $[\phi_i(x)]$ of simple functions. It follows from a theorem of Egoroff's ‡ that this sequence approaches $f(x)$ uniformly on X minus a set of points of arbitrarily small measure. Let $\delta > 0$ be arbitrarily assigned and let a subset E of X be chosen so that E is of measure less than δ and $[\phi_i(x)]$ approaches $f(x)$ uniformly on $X - E$. Let $X_1, X_2, \cdots$ be a countable set of intervals of total length less than δ which covers E and let $M(\delta)$ be the subset of M that belongs to $X - (X_1 + X_2 + \cdots)$. Let $f(x, \delta) = f(x)$ on $M(\delta)$, $f(x, \delta) = 0$ on $X - M(\delta)$. For each $i = 1, 2, 3, \cdots$, let simple functions $g_i(x)$ and $h_i(x)$ be defined as follows: Choose an index n_i so that $| f(x) - \phi_{n_i}(x) | < 1/i$ on $M(\delta)$ and let $g_i(x) = h_i(x) = 0$ on $(X_1 + X_2$

---

* *Bulletin of the American Mathematical Society*, Vol. 37 (1931), page 564.

† *Loc. cit.*, page 561.

‡ See Hobson, *Functions of a Real Variable*, Cambridge Univ. Press, Vol. 2 (1926), p. 140.

$+\cdots+X_i)$; $g_i(x)=\phi_{n_i}(x)-1/i$, $h_i(x)=\phi_{n_i}(x)+1/i$ at all other points of X. Clearly $g_i(x)$ and $h_i(x)$ are simple functions on X as we can make the subdivision points consist of the end-points of the intervals $X_1, X_2, \cdots, X_i$ together with the subdivision points for ϕ_{n_i}. Let $[g_i(x)]$ and $[h_i(x)]$ denote the collections of simple functions obtained for $i=1,2,3,\cdots$, let G_1 and G_2, respectively, denote the collection $[g_i(x)]$ *enlarged below* and the collection $[h_i(x)]$ *enlarged above.* We have $\lim_{i\to\infty} g_i(x) = \lim_{i\to\infty} h_i(x) = f(x,\delta)$ on X_0 and $g_i(x) \leqq f(x,\delta) \leqq h_i(x)$ on X_0 for $i=1,2,3,\cdots$. The separation G_1+G_2 is a cut in K on X that defines $f(x,\delta)$ on X_0. Hence $f(x,\delta)$ is in $C(X)$ and $D(M)$ contains $f(x)$.

Now suppose that M is measurable and $f(x)$ belongs to $D(M)$. Define $f(x)$ to be zero at points of $X-M$ and let i be a positive integer. The function $f(x,1/i)$ belongs to $C(X)$ and, by Theorem V, is measurable on X. Since, by Egoroff's theorem,* $\lim_{i\to\infty} f(x,1/i)=f(x)$ on X_0, it follows immediately that $f(x)$ is measurable on X and hence is measurable on the measurable subset M of X.

THEOREM VIII. *If $f(x)$ belongs to $C(X)$, $\int_a^b f(x)dx$ is equal to the Lebesgue integral of $f(x)$ between $x=a$ and $x=b$.*

Proof. This theorem is an immediate consequence of Theorem V, its corollary, and inequality (1) used in the proof of Theorem V, when account is taken of a theorem of W. M. Whyburn's.†

THEOREM IX. *If H is the subset of M on which $f(x)\neq 0$, where $f(x)$ belongs to $D(M)$, then $\int_{(M)} f(x)dx = \int_{(H)} f(x)dx$ and these integrals are equal to the Lebesgue integral of $f(x)$ on H.*

Proof. Let $f(x)=0$ at points of $X-M$. The sequence of functions $f(x,1/i)$, where $i=1,2,3,\cdots$, approaches $f(x)$ on X_0. If we make use of Theorem VIII and a well-known theorem of Lebesgue's,‡ we get that the limit of the integrals of the functions of this sequence is the Lebesgue integral of $f(x)$ on X. Since $f(x)$ is Lebesgue integrable and bounded on X, it follows that $f(x)$ is measurable on X and hence the set of points H on which $f(x)\neq 0$ is measurable. The Lebesgue integral of $f(x)$ over H is, by definition, equal

* See Hobson, *loc. cit.*, page 140.

† *Bulletin of the American Mathematical Society*, Vol. 38 (1932), page 129, Theorem 7.

‡ *Leçons sur l'intégration etc.*, Paris, 1904, page 114.

to the Lebesgue integral of $f(x)$ on X. The theorem follows immediately from this and the definition of $\int_{(M)} f(x)\,dx$.

COROLLARY. *If $f(x)$ is a function of class $D(M)$, the set of points H is measurable, where H is the subset of M on which $f(x) \neq 0$.*

5. *Remarks.* We have shown how the Dedekind cut notion can be used to give a treatment of the Lebesgue integral of a bounded function. Our work has led exactly to the class of all bounded Lebesgue-integrable functions on an interval X or any measurable subset of X. The integrals of all such functions have been introduced from the Dedekind cut point of view. We might continue the development further to include unbounded, summable functions. The omission of the word bounded from the definition of class $D(M)$ would yield a class which contains all such functions. This further development does not seem necessary, however, since it would consist of an adaptation of the Lebesgue method of building a treatment of integrals of unbounded functions on a previous treatment of integrals of bounded functions. The present treatment uses no more of the theory of measure than is contained in the notions of *δ-set* and *null set*. *Measurable* and *measure of a point set* may be defined through the treatment by saying that a point set M is *measurable* if the function $f(x) = 1$ at points of M, $f(x) = 0$ at points of $X - M$ is in $D(M)$. The *measure* of a *measurable* point set M would then be defined as $\int_a^b f(x)\,dx$, where $f(x) = 1$ on M, $f(x) = 0$ on $X - M$.

THE UNIVERSITY OF CALIFORNIA AT
LOS ANGELES, CALIF.

THE APPLICATION OF BERNOULLI POLYNOMIALS OF NEGATIVE ORDER TO DIFFERENCING.*

By B. F. Kimball.

1. *The expansion of the n-th difference in terms of Bernoulli Functions of Negative Order.* Sometimes it is simpler to calculate the n-th derivative of a function than to calculate the n-th difference. In such cases a series expansion in terms of the derivatives of the function may be of use. We employ the definition and notation for *difference* given by Nörlund.† We use his definition of the Bernoulli function of negative order,‡ which for equal difference intervals becomes

$$(1.1)\quad B_r^{-n}(x, w) = [r!/(n+r)!] \overset{n}{\underset{w}{\Delta}} (x)^{n+r} = w^r[r!/(n+r)!] \overset{n}{\underset{1}{\Delta}} (x/w)^{n+r}.$$

If the difference interval is equal to unity we shall omit the subscript with the operator Δ, and w will be omitted in writing the Bernoulli function. Thus

$$B_r^{-n}(x, w) = w^r B_r^{-n}(x/w) = w^r[r!/(n+r)!]\, \Delta^n (x/w)^{n+r}.$$

If we expand a real analytic function $f(x)$ of real variable x in Taylor's series about a point x_0 and then take the n-th difference of the function and the series term by term, we shall get an expansion formula for the n-th difference of the function in terms of its derivatives at $x = x_0$ and the n-th differences of various powers of $(x - x_0)$. The n-th difference of the term in $(x - x_0)^{n+r}$ is

$$\begin{aligned}[f^{[n+r]}(x_0)/(n+r)!] \overset{n}{\underset{w}{\Delta}} (x - x_0)^{n+r} &= [f^{[n+r]}(x_0)/r!]\, B_r^{-n}(x - x_0, w) \\ &= [w^r f^{[n+r]}(x_0)/r!]\, B_r^{-n}[(x - x_0)/w]\end{aligned}$$

and the expansion formula can be written

$$(1.2)\qquad \overset{n}{\underset{w}{\Delta}} f(x) = \sum_{r=0}^{\infty} (w^r/r!) f^{[n+r]}(x_0) B_r^{-n}\,[(x - x_0)/w].$$

Now it can be shown that $B_{2m+1}^{-n}(-\tfrac{1}{2}n) = 0$ for all positive integral m,

* Presented to the Society, March 25, 1932 [Abstract No. 94, *Bulletin of the American Mathematical Society*, Vol. 38 (1932), p. 186].

† Nörlund, *Differenzenrechnung* (1924), p. 3.

‡ Nörlund, *loc. cit.*, p. 138.

and $m=0$.* Thus formula (1.2) is in many cases simpler if one takes $x-x_0=-\frac{1}{2}nw$. This gives

$$(1.3)\qquad \overset{n}{\underset{w}{\Delta}} f(x)=\sum_{m=0}^{\infty}[w^{2m}/(2m)!]\,B_{2m}^{-n}(-\tfrac{1}{2}n)f^{[n+2m]}(x+\tfrac{1}{2}nw).$$

It will be shown later that *the quantities* $B_{2m}^{-n}(-\frac{1}{2}n)$ *are positive for all positive integral values of* m *and* n [see § 2, formula (2.12)].

This expansion enables one to write $B_r^{-n}(x,w)$ in simple form. Setting $f(x)=[r!/(n+r)!]\,x^{n+r}$, we have from (1.3) (*a*) *when* $r=2m$,

$$(1.4)\qquad B_{2m}^{-n}(x,w)=\sum_{\mu=0}^{m}w^{2\mu}\binom{2m}{2\mu}B_{2\mu}^{-n}(-\tfrac{1}{2}n)\,(x+\tfrac{1}{2}nw)^{2(m-\mu)}$$

where $\binom{2m}{2\mu}$ is the usual symbol for the binomial coefficient; (*b*) *when* $r=2m+1$,

$$(1.5)\qquad B_{2m+1}^{-n}(x,w)=\sum_{\mu=0}^{m}w^{2\mu}\binom{2m+1}{2\mu}B_{2\mu}^{-n}(-\tfrac{1}{2}n)\,(x+\tfrac{1}{2}nw)^{2(m-\mu)+1}.$$

Since $B_{2m}^{-n}(-\frac{1}{2}n)$ is positive for positive n, from (1.4) we judge that *the function* $B_{2m}^{-n}(x,w)$ *is a positive increasing function of* x *when* $x>-\frac{1}{2}nw$ *which becomes infinite as* x *becomes infinite. It takes on its minimum at* $x=-\frac{1}{2}nw$ *and satisfies the relation*

$$(1.6)\qquad B_{2m}^{-n}(-\tfrac{1}{2}nw-x,w)=B_{2m}^{-n}(-\tfrac{1}{2}nw+x,w).$$

From (1.5) we deduce that *the function* $B_{2m+1}^{-n}(x,w)$ *is a positive increasing function of* x *for* $x>-\frac{1}{2}nw$, *becomes infinite as* x *becomes infinite, is equal to zero at* $x=-\frac{1}{2}nw$ *and satisfies the relation*

$$(1.7)\qquad B_{2m+1}^{-n}(-\tfrac{1}{2}nw-x,w)=-B_{2m+1}^{-n}(-\tfrac{1}{2}nw+x,w).$$

2. *Calculation of the function* $B_{2m}^{-n}(-\frac{1}{2}n)$.† From the formulae (1.4)-(1.5) it is seen that the functions $B_{2m}^{-n}(-\frac{1}{2}n)$ characterize the Bernoulli functions of negative order. We have seen that they are of importance in the expansion of a difference in terms of derivatives. A recursion formula for the calculation of $B_{2m}^{-n}(-\frac{1}{2}n)$ can be obtained as follows. Write

$$B_r^{-(n+1)}(x)=\frac{r!}{(n+r+1)!}\,\Delta^{n+1}x^{n+r+1}$$
$$=\frac{1}{r+1}\,\Delta'\frac{(r+1)!}{(n+r+1)!}\,\Delta^n x^{n+r}=\frac{1}{r+1}\,\Delta' B_{r+1}^{-n}(x).$$

* Compare Nörlund, *loc. cit.*, p. 140.

† Compare Nörlund, *loc. cit.*, p. 139, definition of D_{2m}^{-n}.

It can be easily shown that

$$dB_r^{-n}(x)/dx = rB_{r-1}^{-n}(x).$$

Hence expanding $\Delta' B_{r+1}^{-n}(x)$ in series of derivatives, and setting $x = -\frac{1}{2}(n+1)$, we have by (1.3)

$$B_{2m}^{-(n+1)}\left(-\frac{n+1}{2}\right) = \frac{1}{2m+1}\Delta' B_{2m+1}^{-n}\left(-\frac{n+1}{2}\right)$$

$$= \frac{1}{2m+1}\left[(2m+1)B_{2m}^{-n}\left(-\frac{n}{2}\right)\right.$$

$$\left. + (2m+1)\sum_{\mu=1}^{m}\binom{2m}{2\mu}B_{2\mu}^{-1}\left(-\frac{1}{2}\right)B_{2(m-\mu)}^{-n}\left(-\frac{n}{2}\right)\right].$$

Now $B_0^{-n}(-\frac{1}{2}n) = 1$, and it can be easily shown from definition that

$$B_{2m}^{-1}(-\tfrac{1}{2}) = [1/(2m+1)](\tfrac{1}{2})^{2m}.$$

Thus we have

$$(2.1)\qquad B_{2m}^{-(n+1)}\left(-\frac{n+1}{2}\right) = B_{2m}^{-n}(-\tfrac{1}{2}n) + \sum_{\mu=1}^{m}\binom{2m}{2\mu}\frac{1}{2\mu+1}(\tfrac{1}{2})^{2\mu}B_{2(m-\mu)}^{-n}(-\tfrac{1}{2}n).$$

This formula enables one to calculate $B_{2m}^{-n}(-\frac{1}{2}n)$ as a function of n.

On account of the relation

$$(2.2)\qquad \sum_{s=1}^{n-1}\binom{s}{r} = \binom{n}{r+1}$$

it is convenient to express the result in terms of the binomial coefficients $\binom{n}{r}$. For brevity write

$$(2.3)\qquad b_{2k} = B_{2k}^{-1}(-\tfrac{1}{2}) = [1/(2k+1)](\tfrac{1}{2})^{2k}.$$

It is easily shown from formula (2.1) that

$$(2.4)\qquad B_2^{-n}(-\tfrac{1}{2}n) = b_2\cdot\binom{n}{1}.$$

Case 1. *Take* $m = 2$. We have from (2.1) that

$$B_4^{-(n+1)}\left(-\frac{n+1}{2}\right) = B_4^{-n}(-\tfrac{1}{2}n) + \binom{4}{2}b_2B_2^{-n}(-\tfrac{1}{2}n) + b_4$$

since $B_0^{-n}(-\frac{1}{2}n) = 1$.

Now $B_4^{-1}(-2/2) = B_4^{-1}(-\frac{1}{2}) + \binom{4}{2}b_2B_2^{-1}(-\frac{1}{2}) + b_4$

$$= 2b_4 + \binom{4}{2}b_2B_2^{-1}(-\tfrac{1}{2}).$$

Similarly $B_4^{-3}(-3/2) = 3b_4 + \binom{4}{2}b_2[B_2^{-1}(-\frac{1}{2}) + B_2^{-2}(-2/2)]$.

Generalizing, we have

$$(2.5)\qquad B_4^{-n}(-\tfrac{1}{2}n) = nb_4 + \binom{4}{2} b_2 \sum_{s=1}^{n-1} B_2^{-s}(-s/2).$$

Using (2.2) and (2.4) we have

$$(2.6)\qquad \sum_{s=1}^{n-1} B_2^{-s}(-s/2) = b_2 \cdot \binom{n}{2}.$$

Thus

$$(2.7)\qquad B_4^{-n}(-\tfrac{1}{2}n) = \binom{4}{2} b_2^2 \cdot \binom{n}{2} + b_4 \cdot \binom{n}{1}.$$

This may be written

$$(2.8)\qquad B_4^{-n}(-\tfrac{1}{2}n) = \frac{4!}{(2!)^2} b_2^2 \cdot \binom{n}{2} + \frac{4!}{4!} b_4 \cdot \binom{n}{1}$$

and one notes that

$$(2.9)\qquad \sum_{s=1}^{n-1} B_4^{-s}(-\tfrac{1}{2}s) = \frac{4!}{(2!)^2} b_2^2 \cdot \binom{n}{3} + \frac{4!}{4!} b_4 \cdot \binom{n}{2}.$$

Case 2. General Case, m is any positive integer. Reviewing the method of derivation of (2.5) it is not difficult to establish the formula

$$(2.10)\qquad B_{2(m+1)}^{-n}(-\tfrac{1}{2}n) = b_{2(m+1)} \cdot \binom{n}{1} + \sum_{\mu=1}^{m} \binom{2m+2}{2\mu} b_{2\mu} \sum_{s=1}^{n-1} B_{2(m+1-\mu)}^{-s}(-\tfrac{1}{2}s).$$

Calculating $B_{2m}^{-n}(-\tfrac{1}{2}n)$ for several specific values of m [see (2.13)], the following general formula is suggested: Let $p, q, r, \cdots$, be positive integers or zero, and $t, u, v, \cdots$, be positive even integers or zero where

$$pt + qu + rv +, \cdots = 2m$$
$$p + q + r +, \cdots = k.$$

Then

$$(2.11)\qquad B_{2m}^{-n}(-\tfrac{1}{2}n)$$
$$= \sum \frac{k!}{p!\,q!\,r!\cdots} \cdot \frac{(2m)!}{(t!)^p (u!)^q (v!)^r \cdots} \cdot [b_t^p b_u^q b_v^r \cdots] \binom{n}{k}$$

where the summation is extended for all possible values of $p, q, r, \cdots$, and $t, u, v, \cdots$, which yield different combinations of $t^p, u^q, v^r, \cdots$, change of order of these quantities not being considered a change in combination. Reasoning from the assumption that (2.11) is true, using (2.10) it can be shown that (2.11) holds when m is replaced by $(m+1)$. Furthermore (2.11) can easily be checked for small values of m. Thus one concludes that formula (2.11) is true for all values of m. Recalling that $b_t = [1/(t+1)] \times (\tfrac{1}{2})^t$ we have *

* Compare Nörlund, *loc. cit.*, expression for D_{2m}^{-n} on page 140.

(2.12) $B_{2m}^{-n}(-\frac{1}{2}n)$

$$= \frac{1}{2^{2m}} \sum \frac{k!}{p!\,q!\,r!\cdots} \cdot \frac{(2m)!}{[(t+1)!]^p\,[(u+1)!]^q\,[(v+1)!]^r\cdots}\binom{n}{k}$$

the summation being taken as in (2.11).

Note that *if $n \geqq m$, k will take on all integral values from 1 to m inclusive.*

For small values of m we have:

$$B_2^{-n}(-\tfrac{1}{2}n) = \frac{2!}{2^2}\cdot\frac{1}{3!}\binom{n}{1}$$

$$B_4^{-n}(-\tfrac{1}{2}n) = \frac{4!}{2^4}\left[\frac{1}{(3!)^2}\binom{n}{2} + \frac{1}{5!}\binom{n}{1}\right]$$

$$B_6^{-n}(-\tfrac{1}{2}n) = \frac{6!}{2^6}\left[\frac{1}{(3!)^3}\binom{n}{3} + \frac{2}{(3!)(5!)}\binom{n}{2} + \frac{1}{7!}\binom{n}{1}\right]$$

(2.13) $$B_8^{-n}(-\tfrac{1}{2}n) = \left[\frac{8!}{2^8}\,\frac{1}{(3!)^4}\binom{n}{4} + \frac{3}{(3!)^2(5!)}\binom{n}{3} + \left[\frac{1}{(5!)^2} + \frac{2}{(3!)(7!)}\right]\binom{n}{2} + \frac{1}{9!}\binom{n}{1}\right]$$

$$B_{10}^{-n}(-\tfrac{1}{2}n) = \frac{10!}{2^{10}}\left\{\frac{1}{(3!)^5}\binom{n}{5} + \frac{4}{(3!)^3(5!)}\binom{n}{4} + \left[\frac{3}{(3!)(5!)^2} + \frac{3}{(3!)^2\,7!}\right]\binom{n}{3} + \left[\frac{2}{3!\,9!} + \frac{2}{5!\,7!}\right]\binom{n}{2} + \frac{1}{11!}\binom{n}{1}\right\}.$$

Noting that for $n < m$ one or more of the terms $\binom{n}{k}$ are zero one finds that *for $n \geqq m$ the only expressions which vary with n are the functions $\binom{n}{k}$.*

Thus the theorem:

THEOREM 1. *The quantity $B_{2m}^{-n}(-\frac{1}{2}n)$ can be expressed as a linear combination of the functions $\binom{n}{s}$, where s varies from 1 to m, and the coefficients of $\binom{n}{s}$ are positive quantities which are functions of m alone. The explicit formula for $B_{2m}^{-n}(-\frac{1}{2}n)$ is given by* (2.12).

Several deductions can be made from this theorem.

COROLLARY 1.

$$\lim_{n\to\infty} \frac{B_{2m}^{-n}(-\frac{1}{2}n)}{n^m} = \frac{1}{2^{2m}}\cdot\frac{1}{(3!)^m}\cdot\frac{(2m)!}{m!}.$$

This follows at once from formula (2.12). When $x \neq -\frac{1}{2}n$, using formula (1.4) and (1.5) which express $B_r^{-n}(x, w)$ in terms of $B_{2\mu}^{-n}(-\frac{1}{2}n)$ we derive the result:

COROLLARY 2. *The asymptotic value of* $B_r^{-n}(x, w)$ *for* x *constant when* n *becomes infinite is* $(x + \frac{1}{2}nw)^r$.

In discussing the uniform convergence of series (1. 3) for large values of n the following corollary will be found useful.

COROLLARY 3. *For* $n \geqq 3m$ *and* $k \geqq (3/2)m$, *the quotient* $B_{2m}^{-n}(-\frac{1}{2}n)/n^k$ *decreases as* n *increases,* m *and* k *having been held constant.*

In order to establish this corollary we prove the following lemma, and in consequence of this lemma and of Theorem 1; Corollary 3 can be easily deduced.

LEMMA 1. *If* $n \geqq 3m$ *and* $k \geqq (3/2)m$ *then the quotient* $\binom{n}{m}/n^k$ *decreases as* n *increases,* k *and* m *being held constant.*

To prove this lemma note that

$$\frac{d\binom{n}{m}}{dn} = \binom{n}{m}\left[\frac{1}{n} + \frac{1}{n-1} + \cdots \frac{1}{n-(m-1)}\right] < \binom{n}{m}\log\frac{n}{n-m}.$$

Hence
$$\frac{d[\binom{n}{m}/n^k]}{dn} < \frac{\binom{n}{m}\left[n\log\dfrac{n-m}{n} - k\right]}{n^{k+1}}.$$

Hence if $k > n \cdot \log[n/(n-m)]$, the derivative will be negative.

Set $n = hm$ and the condition becomes

$$k > m\left(h\log\frac{h}{h-1}\right).$$

Now if $\qquad h \geqq 3, \qquad h\log[h/(h-1)] < 3/2.$

Thus the lemma follows easily.

3. *An upper bound of the function* $B_{2m}^{-n}(-\frac{1}{2}n)$. A recursion formula obtained by Nörlund * enables one to establish an upper bound to $B_{2m}^{-n}(-\frac{1}{2}n)$ which is useful in studying series expansions in terms of these functions. For Bernoulli functions of negative order, it is

$$B_r^{-(n-1)}(x) = (1 + r/n)B_r^{-n}(x) - (x+n)(r/n)B_{r-1}^{-n}(x)$$

and solving for $B_r^{-n}(x)$ we have

$$(3.1) \qquad B_r^{-n}(x) = \frac{n}{n+r}B_r^{-(n-1)}(x) + \frac{(x+n)r}{n+r}B_{r-1}^{-n}(x).$$

Take $r = 2m$, $x = -\frac{1}{2}n$. Then since $B_{2m-1}^{-n}(-\frac{1}{2}n) = 0$, we have

* Nörlund, *loc. cit.*, formula (81), p. 145.

$$(3.2) \qquad B_{2m}^{-n}(-\tfrac{1}{2}n) = \frac{n}{n+2m} B_{2m}^{-(n-1)}(-\tfrac{1}{2}n).$$

Again, write $B_{2m}^{-(n-1)}(-\tfrac{1}{2}n)$ by means of (3. 1) in the form

$$B_{2m}^{-(n-1)}(-\tfrac{1}{2}n) = \frac{n-1}{n-1+2m} B_{2m}^{-(n-2)}(-\tfrac{1}{2}n) + \frac{(\tfrac{1}{2}n-1)\cdot 2m}{n-1+2m} B_{2m-1}^{-(n-1)}(-\tfrac{1}{2}n).$$

Since $-\tfrac{1}{2}n < [-(n-1)/2]$, it follows that $B_{2m-1}^{-(n-1)}(-\tfrac{1}{2}n)$ is negative [see § 1 (1. 7)]. Thus

$$(3.3) \qquad B_{2m}^{-(n-1)}(-\tfrac{1}{2}n) < \frac{n-1}{n-1+2m} B_{2m}^{-(n-2)}(-\tfrac{1}{2}n).$$

Replacing n in (3. 1) by $n-2$, we have

$$B_{2m}^{-(n-2)}(-\tfrac{1}{2}n) = \frac{n-2}{n-2+2m} B_{2m}^{-(n-3)}(-\tfrac{1}{2}n) + \frac{(\tfrac{1}{2}n-2)\cdot 2m}{n-2+2m} B_{2m-1}^{-(n-2)}(-\tfrac{1}{2}n).$$

Again $B_{2m-1}^{-(n-2)}(-\tfrac{1}{2}n)$ is negative and hence

$$(3.4) \qquad B_{2m}^{-(n-2)}(-\tfrac{1}{2}n) < \frac{n-2}{n-1+2m} B_{2m}^{-(n-3)}(-\tfrac{1}{2}n).$$

Similar inequalities to (3. 3) and (3. 4) are obtained when the order of function on left is $n-k$ provided $n-k > \tfrac{1}{2}n$. If $n-k = \tfrac{1}{2}n$, we have

$$(3.5) \qquad B_{2m}^{-n/2}(-\tfrac{1}{2}n) = \frac{\tfrac{1}{2}n}{\tfrac{1}{2}n+2m} B_{2m}^{-(n/2-1)}(-\tfrac{1}{2}n) + 0.$$

Combining the above results we have for n even

$$(3.6) \quad B_{2m}^{-n}(-\tfrac{1}{2}n) < \frac{n(n-1)(n-2)\cdots(\tfrac{1}{2}n)}{(n+2m)(n-1+2m)\cdots(\tfrac{1}{2}n+2m)} B_{2m}^{-(n/2-1)}(-\tfrac{1}{2}n), \quad n > 2.$$

For n odd, letting $[\tfrac{1}{2}n]$ denote greatest integer in $\tfrac{1}{2}n$, we have

$$(3.7) \quad B_{2m}^{-n}(-\tfrac{1}{2}n) < \frac{n(n-1)(n-2)\cdots([\tfrac{1}{2}n]+1)}{(n+2m)(n-1+2m)\cdots([\tfrac{1}{2}n]+1+2m)} B_{2m}^{-[n/2]}(-\tfrac{1}{2}n), \quad n > 1.$$

Now it is well known * that

$$\Delta^k f(x) = f^{[k]}(x+u), \qquad 0 < u < k.$$

Thus from the definition of $B_{2m}^{-k}(x)$ and this equation it follows that

$$B_{2m}^{-k}(x) = (x+v)^{2m}, \qquad 0 < v < k,$$

* Nörlund, *loc. cit.*, formula (29), p. 13.

7

and accordingly that

$$B_{2m}^{-(n/2-1)}(-\tfrac{1}{2}n) = (-\tfrac{1}{2}n + v_1)^{2m}, \qquad 0 < v_1 < (\tfrac{1}{2}n - 1),$$
$$B_{2m}^{-[n/2]}(-\tfrac{1}{2}n) = (-\tfrac{1}{2}n + v_2)^{2m}, \qquad 0 < v_2 < [\tfrac{1}{2}n].$$

Thus $(\frac{1}{2}n)^{2m}$ is an upper bound to both of these last functions. Hence the theorem:

THEOREM 2. *For all positive integral values of m and n the function $B_{2m}^{-n}(-\frac{1}{2}n)$ satisfies the relation*

$$(3.8) \qquad B_{2m}^{-n}(-\tfrac{1}{2}n) \leqq \frac{n(n-1)\cdots r}{(n+2m)(n-1+2m)\cdots(r+2m)} \cdot (\tfrac{1}{2}n)^{2m}$$

where $r = [(n+1)/2]$. The equal sign applies when $n \leqq 2$ and only then.

A consequence of this theorem is the corollary:

COROLLARY 1. *The series*

$$\sum_{m=1}^{\infty} B_{2m}^{-n}(-\tfrac{1}{2}n)/(\tfrac{1}{2}n)^{2m}$$

will converge when $n > 1$. It diverges for $n = 1$.

4. *Asymptotic value of the function of n represented by*

$$\sum_{m=1}^{\infty} [B_{2m}^{-n}(-\tfrac{1}{2}n)/(\tfrac{1}{2}n)^{2m}].$$

Since this series occurs quite often in differencing, the following theorem is of interest.

THEOREM 3. *Let the quotient $B_{2m}^{-n}(-\frac{1}{2}n)/(\frac{1}{2}n)^{2m}$ be denoted by $w(m,n)$. Given k, any positive integer, the asymptotic value of $n^k \sum_{m=1}^{\infty} w(m,n)$ is the same as that of $n^k \sum_{m=1}^{k} w(m,n)$.*

Since $B_{2m}^{-n}(-\frac{1}{2}n)$ for $n > m$ is of the m-th degree in n we see that $n^k \sum_{m=1}^{k} w(m,n)$ is at least of degree zero in n and that it approaches a positive constant or becomes infinite as n becomes infinite. Thus, *in order to prove the above theorem it will be sufficient to show that $n^k \sum_{m=k+1}^{\infty} w(m,n)$ approaches zero as n becomes infinite.* To that end we break this series up into the three parts:

$$(4.1) \qquad P(n) = n^k \sum_{m=k+1}^{2k} w(m,n), \qquad Q(n) = n^k \sum_{m=2k+1}^{[n/3]} w(m,n),$$
$$R(n) = n^k \sum_{m=[n/3]+1}^{\infty} w(m,n).$$

In the case of the first part we deal with a finite number of terms and it follows at once from Corollary 1, Theorem 1 that $P(n)$ approaches zero as n becomes infinite.

In the case of the second part $Q(n)$ we find that the hypotheses of Corollary 3, Theorem 1 are satisfied; namely, $n \geqq 3m$ and the exponent $2m - k > (3/2)m$ for all admissable m and n. Thus, *each term* $n^k w(m, n)$ *of* $Q(n)$ *decreases as n increases.* Since, however, new terms are added when n increases we shall find it convenient to employ the majorante series formed by taking $n = 3m$ in each term of the original series. We shall have

$$(4.2) \qquad n^k \sum_{m=2k+1}^{[n/3]} w(m, n) < \sum_{m=2k+1}^{[n/3]} (3m)^k w(m, 3m),$$

since for every term $n^k w(m, n)$ on the left, except possibly the last term $m = [n/3]$, $n > 3m$ and therefore by Corollary 3, Theorem 1 this term is less than the corresponding term formed by setting $n = 3m$. If $n/3$ is an integer, the last term $m = [n/3]$ is *equal* to the corresponding term of the series on the right, otherwise it is less than that term. The inequality (4.2) is thus true for all large n.

In order to establish the convergence of the new series as n becomes infinite recall that from Theorem 2, it follows that

$$(4.3) \qquad w(m, n) < u(m, n) = \frac{n(n-1)\cdots r}{(n+2m)(n-1+2m)\cdots(r+2m)}, \quad n > 2,$$

where $r = [(n+1)/2]$. Thus the above majorante series as n becomes infinite is less, term by term, than the series

$$(4.4) \qquad \sum_{m=k+1}^{\infty} (3m)^k u(m, 3m).$$

The test ratio for convergence of (4.4) is easily shown to approach a limit less than unity. *Thus the majorante series of* (4.2) *converges as n becomes infinite.*

It follows from the following argument that

$$(4.5) \qquad \lim_{n\to\infty} Q(n) = 0.$$

Take m_0 large enough so that for positive ϵ chosen in advance,

$$(4.6) \qquad \sum_{m=m_0}^{\infty} (3m)^k w(m, 3m) < \epsilon.$$

Then for $n > 3m_0$ we can write

$$Q(n) = n^k \sum_{m=2k+1}^{m_0} w(m, n) + n^k \sum_{m=m_0}^{[n/3]} w(m, n).$$

The second expression on the right is always less than ϵ by (4. 2) and (4. 6). The first expression on the right involves only a finite number of terms and since each term approaches zero as $n \to \infty$ the sum approaches zero as $n \to \infty$. Thus the truth of (4. 5) is demonstrated.

We have left to show that

$$\lim_{n\to\infty} R(n) = 0. \tag{4. 7}$$

Write $R(n)$ in the form

$$R(n) = n^k \sum_{t=1}^{\infty} w(q+t, n), \qquad q = [n/3]. \tag{4. 8}$$

It is found convenient to introduce the following lemma:

LEMMA 2.

$$\lim_{n\to\infty} [\sum_{t=1}^{\infty} 2^{n/3} w(q+t, n)] = 0, \quad \text{where} \quad q = [n/3]. \tag{4. 9}$$

Using $u(m, n)$ as defined in (4. 3),

$$\sum_{t=1}^{\infty} 2^{n/3} w(q+t, n) < < \sum_{t=1}^{\infty} 2^{n/3} u(q+t, n), \qquad q = [n/3]. \tag{4. 10}$$

The majorante series on the right is known to converge for fixed n greater than one.

We shall show that it converges uniformly for large values of n. In order to do this first consider the case $n \equiv 0$ *mod* 6, and study the change in the t-th term when n is increased by 6. The ratio of the new term to the old will be given by

$$\rho = 4 \frac{u[(n+6)/3 + t, n+6]}{u(n/3 + t, n)} . \tag{4. 11}$$

A simple calculation shows that the limit of this ratio as n becomes infinite is $\frac{7^7 \; 3^3}{5^{10} \; 2^2} = .57^{-}$. It is also easily shown that for any fixed n, ρ decreases as t increases. We thus judge that it is possible to find n_0 for given t_0 such that $\rho < 1$ when $n \geqq n_0$ uniformly in t for $t \geqq t_0$. Hence for $n \geqq n_0$ all the terms of the majorante series of (4. 10) decrease when n increases by six. This we have shown to be true for the case $n \equiv 0$ *mod* 6. A similar argument will lead to the same conclusion for the cases $n \equiv 1, 2, \cdots, 5$ *mod* 6. It follows that *the majorante series of* (4. 10) *is uniformly convergent* $n \geqq n_0$.

Now it can be easily shown that for every value of t

$$\lim_{n\to\infty} 2^{n/3} u(q+t, n) = 0, \qquad q = [n/3].$$

Hence

$$\lim_{n\to\infty} \sum_{t=1}^{\infty} 2^{n/3} u(q+t, n) = 0, \qquad q = [n/3],$$

and using (4.10) the proof of Lemma 2 is complete.

If we write

$$\lim n^k \sum_{t=1}^{\infty} w(q+t,n) = \lim [n^k \cdot (\tfrac{1}{2})^{n/3}] \cdot \lim \sum_{t=1}^{\infty} 2^{n/3} w(q+t,n)$$

it is clear in the light of Lemma 2 that this limit is zero as $n \to \infty$. Thus (4.7) is demonstrated and Theorem 3 is established.

Recalling that $B_2^{-n}(-\frac{1}{2}n) = n/12$ [v. (2.13)], we have $w(1,n) = 1/3n$ and thus the case $k = 1$ of Theorem 3 leads to the corollary:

Corollary 1. *The asymptotic value of* $n \sum_{m=1}^{\infty} w(m,n)$ *as* $n \to \infty$ *is* 1/3.

The content of Theorem 3 can be extended considerably by the corollary

Corollary 2. *Given* k *any positive integer, and the series*

$$n^k \sum_{m=1}^{\infty} p(m) w(m,n), \tag{4.12}$$

where $w(m,n)$ *is defined as in Theorem* 3 *and* $p(m)$ *is a real function of* m, *independent of* n, *which takes on positive values for all positive integral* m. *Also given that* $\sum_{m=1}^{\infty} p(m) u(m,n)$ *converges for some value of* n, *say* n_0, *where* $u(m,n)$ *is defined as in* (4.3). *Then the series* (4.12) *converges for* $n \geqq n_0$ *and becomes asymptotic like* $n^k \sum_{m=1}^{k} p(m) w(m,n)$.

That the series (4.12) converges for $n \geqq n_0$ follows from a brief consideration of the test ratio of the majorante series. It is easily shown that this ratio decreases when n is increased by 1. Thus the majorante series will converge $n \geqq n_0$ and *a fortiori* the series (4.12).

In order to discuss the asymptotic behavior of (4.12) we proceed as in the proof of Theorem 3 and write

$$P_1(n) = n^k \sum_{m=k+1}^{2k} p(m)\, w(m,n), \qquad Q_1(n) = n^k \sum_{m=2k+1}^{[n/3]} p(m)\, w(m,n),$$

$$R_1(n) = n^k \sum_{m=[n/3]+1}^{\infty} p(m)\, w(m,n).$$

As in case $P(n)$, $P_1(n) \to 0$ as $n \to \infty$. In the case of $Q_1(n)$, the majorante series used [cf. (4.4)] will be

$$\sum_{m=2k+1}^{\infty} (3m)^k\, p(m)\, u(m, 3m). \tag{4.13}$$

To establish convergence, consider the series $\sum_{m=2k+1}^{\infty} p(m)\, u(m,n)$ where the

summation is for odd values of m, only. This series will converge for $n = n_0$ since it is a part of a convergent series of positive terms. The test ratio satisfies the inequality

$$(4.14) \qquad \frac{p(m+2)\,u(m+2,n)}{p(m)\,u(m,n)} < \frac{p(m+2)\,u(m+2,n_0)}{p(m)\,u(m,n_0)}, \quad n > n_0.$$

Now take the series of terms from (4.13) for which m is odd. Using (4.14) it is not a difficult matter to show that the test ratio of the new series becomes and remains less than the test ratios (4.14) when m becomes infinite. A similar argument will apply when all values of m are taken as even. We conclude that (4.13) is convergent. Proceeding as in the previous case it follows that

$$\lim_{n\to\infty} Q_1(n) = 0.$$

In discussing $R_1(n)$ one shows that

$$(4.15) \qquad \lim_{n\to\infty} \Big[\sum_{t=1}^{\infty} 2^{n/3}\,p(m)\,w(q+t,n)\Big] = 0, \qquad q = [n/3],$$

[cf. Lemma 2, (4.9)]. Since $p(m)$ is independent of n, the ratio ρ for the majorante series $\sum_{t=1}^{\infty} 2^{n/3}\,p(m)\,u(q+t,n)$ will be the same as given by (4.11). Hence the proof of the uniform convergence of this series proceeds as in the previous case. Furthermore the proof of (4.15) and the remainder of the proof that

$$\lim_{n\to\infty} R_1(n) = 0$$

are almost identical with the line of argument pursued in the discussion of $R(n)$. This establishes Corollary 2.

5. *Miscellaneous applications.* (a) *A question of convergence.* The difference expansion series (1.3) may be thought of as obtained from the series:

$$f(x) = \sum_{s=0}^{\infty} \frac{(-\frac{1}{2}n\cdot w)^s}{s!}\, f^{[s]}(x + \tfrac{1}{2}n\cdot w)$$

$$f(x+1) = \sum_{s=0}^{\infty} \frac{(-\frac{1}{2}nw + w)^s}{s!}\, f^{[s]}(x + \tfrac{1}{2}nw)$$

$$\cdots\cdots\cdots\cdots$$

$$f(x+n) = \sum_{s=0}^{\infty} \frac{(+\frac{1}{2}nw)^s}{s!}\, f^{[s]}(x + \tfrac{1}{2}nw).$$

Thus we have the lemma:

LEMMA 3. *The convergence of the series*

$$\sum_{s=n}^{\infty} (z^s/s!) f^s(x + \tfrac{1}{2}nw),$$

for z on the interval $-\tfrac{1}{2}nw \leqq z \leqq + \tfrac{1}{2}nw$ is a sufficient condition for the convergence of the series (1.3).

For example consider the expansion of the n-th difference of $f(x) = x^p$ where p is a non-integral real number. It is not easy to establish the convergence at $x = 0$ directly from this series. However, if we examine the series

$$(5.2) \qquad \sum_{s=0}^{\infty} \binom{p}{s} z^s (\tfrac{1}{2}n)^{p-s} = (\tfrac{1}{2}n)^p \sum_{s=0}^{\infty} \binom{p}{s} (z/\tfrac{1}{2}n)^s$$

over the interval $-\tfrac{1}{2}n \leqq z \leqq + \tfrac{1}{2}n$, we see that it converges throughout the interval in question if $p > 0$. Hence using Lemma 3 one concludes that where p is a non-integral positive number, the expansion of $\Delta^n x^p$ at $x = 0$ of form (1.3) for $w = 1$, converges.

(b) *Consider $f(x) = x^{n+q}$ where n is a positive integer.*

Case I. If q is an integer and $n + q > 0$, Corollary 2, Theorem 1 tells the story about $\Delta^n f(x)$.

Case II. q non-integral and $n + q > 0$. Then using (1.3) we can write $\Delta^n f(x)$ in the form

$$(5.3) \qquad \Delta^n f(x) = f^n(x + \tfrac{1}{2}n) \left[1 + \sum_{m=1}^{\infty} \binom{q}{2m} \cdot \left(\frac{n}{2x + n} \right)^{2m} w(m, n) \right]$$

where $w(m, n)$ is defined as in Theorem 3. We have seen from the example given under Lemma 3 above, that (5.3) converges when $x = 0$. Moreover an investigation of the test ratio of the series $\sum_{m=1}^{\infty} \binom{q}{2m} u(m, n)$ shows it to be convergent certainly for positive n greater than $-2q$. Thus the hypotheses of Corollary 2, Theorem 3 are satisfied with respect to the series $\sum_{m=1}^{\infty} \binom{q}{2m} w(m, n)$. This corollary indicates that

$$\lim_{n \to \infty} n^k \sum_{m=k+1}^{\infty} \binom{q}{2m} w(m, n) = 0.$$

Clearly it follows that

$$(5.4) \qquad \lim_{n \to \infty} n^k \sum_{m=k+1}^{\infty} \binom{q}{2m} \left(\frac{n}{2x + n} \right)^{2m} w(m, n) = 0$$

for positive x. We thus conclude that *in Case II for x positive or zero*

$$(5.5) \qquad \lim_{n \to \infty} \frac{\Delta^n f(x)}{f^{[n]}(x + \tfrac{1}{2}n)} = 1,$$

and $n^k\left[\frac{\Delta^n f(x)}{f^{[n]}(x+\frac{1}{2}n)}-1\right]$ *becomes asymptotic like*

$$n^k\sum_{m=1}^{k}\binom{q}{2m}\left(\frac{n}{2x+n}\right)^{2m}w(m,n).$$

Case III. q *has value such that* $n+q<0$. The expansion of $\Delta^n f(x)$ has the form (5.3). The series converges for positive x but diverges when $x=0$. Since q is fixed there is no asymptotic form of $\Delta^n f(x)$ under Case III.

(c) *Consider* $f(x)=x^{n-q}\log x$, where q is a positive integer less than the positive integer n, and $f(0)$ is defined to be zero. We have:

$$f^{[n]}(x)=(-1)^{q-1}\frac{(n-q)!\,(q-1)!}{x^q},$$

$$f^{[n+2m]}(x)=(-1)^{q-1}\frac{(n-q)!\,(q+2m-1)!}{x^{q+2m}}.$$

Thus using (1.3) we can write:

$$\Delta^n f(x)=f^{[n]}(x+\tfrac{1}{2}n)\left[1+\sum_{m=1}^{\infty}\binom{q+2m-1}{2m}\left(\frac{n}{2x+n}\right)^{2m}w(m,n)\right]. \tag{5.6}$$

Using Lemma 3 it is not difficult to show that this series is convergent for $x=0$. Now the series $\sum_{m=1}^{\infty}\binom{q+2m-1}{2m}u(m,n)$ can be shown to converge for $n>2q$. Thus from Corollary 2, Theorem 3 we infer as in example (b), (5.4), that

$$\lim_{n\to\infty} n^k\sum_{m=k+1}^{\infty}\binom{q+2m-1}{2m}\left(\frac{n}{2x+n}\right)^{2m}w(m,n)=0, \tag{5.7}$$

for x positive or zero. The question as to whether or not the convergent series (5.6) represents correctly $\Delta^n f(0)$ when $x=0$ can be answered as follows. The fact that $f(x)$ is continuous $0\leqq x$ means that $\Delta^n f(x)$ is continuous on this interval. The series (5.6) is a uniformly convergent series of continuous functions with respect to x and hence at $x=0$ sums to $\Delta^n f(0)$. We can thus state that *for x positive or zero,*

$$\lim_{n\to\infty}\frac{\Delta^n f(x)}{f^{[n]}(x+\frac{1}{2}n)}=1 \tag{5.8}$$

and $n^k\cdot\left[\frac{\Delta^n f(x)}{f^{[n]}(x+\frac{1}{2}n)}-1\right]$ *becomes asymptotic like*

$$n^k\cdot\sum_{m=1}^{k}\binom{q+2m-1}{2m}\left(\frac{n}{2x+n}\right)^{2m}w(m,n).$$

It is of interest to note that the series expansions within the brackets of formulae (5.3) and (5.6) become identical when the q used in (5.3) is

the negative of the integer q employed above. The uniform convergence of the series of (5.3) with respect to q neighboring a negative integer (numerically less than n) can easily be established; hence *if* $\phi(x) = x^{n-s}$,

$$(5.9)\qquad \lim_{s\to q}\frac{\Delta^n\phi(x)}{\phi^{[n]}(x+\tfrac{1}{2}n)} = \frac{\Delta^n f(x)}{f^{[n]}(x+\tfrac{1}{2}n)}$$

where $f(x)$ *and* q *are defined as in* (5.6).

(d) Consider the n-fold integral

$$(5.10)\qquad Q(n) = \int_0^1\int_0^1\cdots\int_0^1 (x_1+x_2+\cdots+x_n)^p\,dx_1\,dx_2\cdots dx_n$$

where p is any real quantity. Let $f_n(x)$ denote a function of s such that

$$\frac{d^n f_n(s)}{ds} = s^p, \qquad 0 < s.$$

Now it is well known * that for an improper integral of the above type,

$$(5.11)\qquad Q(n) = \lim_{s\to 0^+}\Delta^n f_n(s).$$

If p is a negative integer numerically less than or equal to n, we can take

$$f_n(s) = (-1)^{q-1}\frac{1}{(n-q)!\,(q-1)!}\,s^{n-q}\log s, \quad q = -p, \quad s > 0;$$
$$f_n(0) = 0.$$

For other values of p, we can set

$$f_n(s) = \frac{1}{(q+1)(q+2)\cdots(q+n)}\cdot s^{n+q}, \qquad q = p.$$

After a brief consideration of the function $f_n(s)$ and (5.11) it is clear that the integral $Q(n)$ converges and is represented by $\Delta^n f_n(0)$ when $n+p>0$; and that *it diverges when* $n+p\leqq 0$. Thus combining (5.3) and (5.6) as indicated by (5.9) we can state that

$$(5.12)\qquad Q(n) = (\tfrac{1}{2}n)^p\left[1+\sum_{m=1}^{\infty}\binom{p}{2m}w(m,n)\right], \qquad n+p>0.$$

and $n^k\left[\dfrac{Q(n)}{(\frac{1}{2}n)^p}-1\right]$ *becomes asymptotic like* $n^k\sum_{m=1}^{k}\binom{p}{2m}w(m,n)$.

6. *Asymptotic value of* $\Delta^n\log x$. Another application is the determination of the asymptotic value of $\Delta^n\log x$ for n infinite. Take $f(x)=1/x$ and

$$(6.1)\qquad H_n(x) = (-1)^n\Delta^n f(x) = \frac{n!}{x(x+1)\cdots(x+n)}\cdot$$

---

* Nörlund, *loc. cit.*, page 14, formula (30).

We consider the case where x is positive and for the sake of definiteness take n as odd. The analysis will follow through equally well when n is even. Then

$$\Delta^n f'(x) = [\Delta^n f(x)]' = (-1)H'_n(x) = H_n(x) \cdot \left(\frac{1}{x} + \frac{1}{x+1} + \cdots \frac{1}{x+n}\right),$$

or

$$\Delta^n f'(x) = -H'_n(x) = H_n(x) u_1(x), \tag{6.2}$$

where

$$u_1(x) = \sum_{i=0}^{n} \frac{1}{x+i}\,.$$

We also set

$$u_s(x) = \sum_{i=0}^{n} \frac{1}{(x+i)^s}\,.$$

Then

$$\Delta^n f''(x) = -H_n''(x) = -H_n(x)[u_1^2(x) + u_2(x)].$$

Similarly

$$\Delta^n f'''(x) = -H_n'''(x) = H_n(x)[(u_1^3(x) + 3u_1(x)u_2(x) + 2u_3(x)],$$

and it is easily shown by induction that

$$\Delta^n f^{[m]}(x) = (-1)^{m+1} H_n(x)[u_1^m(x) + R_m(x)] \tag{6.3}$$

where $R_m(x)$ is made up of only positive terms (x positive). It also can be shown that $R'_m(x)$ possesses only negative terms and $R_m''(x)$ only positive terms. We shall need the following lemma:

LEMMA 4. *For positive x and all positive integral values of m and n*

$$|R'_m(x)| < |R'_{m+1}(x)/u_1(x)| \tag{6.4}$$

where $R_m(x)$ is defined by (6.3).

Calculation of $\Delta^n f^{m+1}(x)$ from (6.3) gives

$$\Delta^n f^{m+1}(x) \\ = (-1)^m H_n(x)[u_1^{m+1}(x) + m u_1^{m-1}(x) u_2(x) + u_1(x) R_m(x) - R'_m(x)].$$

Thus

$$R_{m+1}(x) = m u_1^{m-1}(x) u_2(x) + u_1(x) R_m(x) - R'_m(x). \tag{6.5}$$

Leaving out the indication of the independent variable in the notation we differentiate and obtain

$$R'_{m+1} = -\, m(m-1) u_1^{m-2} u_2^2 - 2m u_1^{m-1} u_3 - u_2 R_m + u_1 R'_m - R_m''.$$

Now all the terms in the last equation are negative. Thus

$$| R'_{m+1}/u_1 | = | R'_m | + R_m''/u_1 + (u_2 R_m)/u_1 + m(m-1)u_1^{m-3}u_2 + 2mu_1^{m-2}u_3$$

and the lemma follows.

We can now attack directly the problem of the determination of the asymptotic value of $\Delta^n \log x$. Consider $\Delta^n \log x$ as $\Delta'(\Delta^{n-1} \log x)$ and expand in terms of first derivatives of $\Delta^{n-1} \log x$ at $x + \frac{1}{2}$ [v. § 1 (1. 3)].

$$(6.6) \quad \Delta^n \log x = \Delta^{n-1} f(x + \tfrac{1}{2}) + \sum_{\mu=1}^{\infty} (1/2\mu!) B_{2\mu}^{-1}(-\tfrac{1}{2}) \Delta^{n-1} f^{2\mu}(x + \tfrac{1}{2}).$$

It is to be noted that all the terms in the above expansion are positive. Similarly since all the odd derivatives of $\Delta^{n-1} f(x + \frac{1}{2})$ are negative we write

$$(6.7) \quad | \Delta^n f(x) | = | \Delta^{n-1} f'(x + \tfrac{1}{2}) | + \sum_{\mu=1}^{\infty} \frac{1}{2\mu!} B_{2\mu}^{-1}(-\tfrac{1}{2}) | \Delta^{n-1} f^{2\mu+1}(x + \tfrac{1}{2}) |$$

and again

$$(6.8) \quad \Delta^n f'(x) = \Delta^{n-1} f''(x + \tfrac{1}{2}) + \sum_{\mu=1}^{\infty} \frac{1}{2\mu!} B_{2\mu}^{-1}(-\tfrac{1}{2}) \Delta^{n-1} f^{2\mu+2}(x + \tfrac{1}{2}).$$

Now values of $| \Delta^n f(x) |$ and $\Delta^n f'(x)$ are known to be $H_n(x)$ and $H_n(x) u_1(x)$ respectively. The method of the determination of $\Delta^n \log x$ will be to establish the double inequality

$$(6.9) \quad 0 < \left| \frac{\Delta^n f(x)}{U_1(x + \frac{1}{2})} \right| - \Delta^n \log x < \frac{1}{U_1(x + \frac{1}{2})} \left[\frac{\Delta^n f'(x)}{U_1(x + \frac{1}{2})} - | \Delta^n f(x) | \right]$$

by means of the above series developments where $U_1(x + \frac{1}{2})$ denotes $u_1(x + \frac{1}{2})$ with n replaced by $n - 1$. The $m + 1$-th term of (6. 6) is

$$(6.10) \qquad \frac{1}{2m!} B_{2m}^{-1}(-\tfrac{1}{2}) H_{n-1}(x + \tfrac{1}{2}) [U_1^{2m} + R_{2m}(x + \tfrac{1}{2})]$$

where U_1 denotes $U_1(x + \frac{1}{2})$. The $m + 1$-th term of (6. 7) is

$$(6.11) \quad \frac{1}{2m!} B_{2m}^{-1}(-\tfrac{1}{2}) H_{n-1}(x + \tfrac{1}{2}) \times [U_1^{2m+1} + U_1 R_{2m}(x + \tfrac{1}{2}) + 2mU_1^{2m-1} U_2 + | R'_{2m}(x + \tfrac{1}{2}) |]$$

where U_2 means $u_2(x + \frac{1}{2})$ with n replaced by $n - 1$. Divide the term (6. 11) by U_1 and subtract the term (6. 10). This yields

$$(6.12) \quad \frac{1}{2m!} B_{2m}^{-1}(-\tfrac{1}{2}) H_{n-1}(x + \tfrac{1}{2}) [| R'_{2m}(x + \tfrac{1}{2}) | / U_1 + 2mU_1^{2m-2} U_2].$$

This is positive for all positive x and for all positive integral m. Hence, noting a similar relation between the first terms of expansions (6. 6) and (6. 7), the left hand side of double inequality (6. 9) is established.

Similarly one finds that the algebraic sum of the $m + 1$-th terms of the series for $\Delta^n f'(x)/U_1(x + \frac{1}{2})$ and $- | \Delta^n f(x) |$ is

$$(6.13) \quad \frac{1}{2m!} B_{2m}^{-1}(-\tfrac{1}{2}) H_{n-1}(x + \tfrac{1}{2}) [| R'_{2m+1}(x + \tfrac{1}{2}) | / U_1 + (2m + 1) U_1^{2m-1} U_2].$$

Hence dividing (6.13) by U_1 and subtracting (6.12) from the quotient we have

$$\frac{1}{2m!} B_{2m}^{-1}(-\tfrac{1}{2}) H_{n-1}(x+\tfrac{1}{2}) \times \left[\frac{|R'_{2m+1}(x+\tfrac{1}{2})|}{U_1^2} - \frac{|R'_{2m}(x+\tfrac{1}{2})|}{U_1} + U_1^{2m-2} U_2 \right].$$

Lemma 4 can now be applied replacing n by $n-1$ and x by $x+\frac{1}{2}$. We find that this expression is positive. Since this is true for all positive integral values of m and since a similar relation holds between the *first* terms of the series (6.7) and (6.8), the double inequality (6.9) follows as a consequence.

The substitution of $H_n(x)$ and $u_1(x)$ in (6.9) gives

$$(6.14) \qquad 0 < \frac{H_n(x)}{U_1(x+\frac{1}{2})} - \Delta^n \log x < \frac{H_n(x)}{U_1(x+\frac{1}{2})} \left[\frac{u_1(x)}{U_1(x+\frac{1}{2})} - 1 \right].$$

Thus

$$(6.15) \qquad 0 < 1 - \frac{U_1(x+\frac{1}{2})}{H_n(x)} \Delta^n \log x < \left[\frac{u_1(x)}{U_1(x+\frac{1}{2})} - 1 \right].$$

Now

$$\lim_{n\to\infty} \frac{u_1(x)}{U_1(x+\frac{1}{2})} = 1, \qquad \lim_{n\to\infty} \frac{U_1(x+\frac{1}{2})}{\log n} = 1, \qquad \lim_{n\to\infty} H_n(x) n^x = \Gamma(x).$$

Hence

$$(6.16) \qquad \lim_{n\to\infty} (n^x \log n)(\Delta^n \log x) = \Gamma(x).$$

Furthermore we can write

$$\Delta^n \log x = \frac{(1-\epsilon) n!}{x(x+1)\cdots(x+n)} \cdot \frac{1}{U_1(x+\frac{1}{2})}$$

$$0 < \epsilon < \frac{u_1(x)}{U_1(x+\frac{1}{2})} - 1.$$

where

$$u_1(x) = \sum_{i=0}^{n} \frac{1}{x+i}, \qquad U_1(x+\tfrac{1}{2}) = \sum_{i=0}^{n-1} \frac{1}{x+\frac{1}{2}+i}.$$

An easy calculation shows that

$$0 < \epsilon < \left| \frac{2x+3}{4x^2} + \frac{1}{2(x+n)} \right] / \log\left(\frac{x+n+\frac{1}{2}}{x+\frac{1}{2}} \right).$$

The above results can be extended to the case where the difference interval is a positive real quantity w by using the relation

$$\underset{w}{\overset{n}{\Delta}} \log(x) = (1/w^n) \underset{1}{\overset{n}{\Delta}} \log(x/w).$$

SCHENECTADY, N. Y.

GROUPS INVOLVING ONLY OPERATORS WHOSE ORDERS DIVIDE 4 AND WHOSE OPERATORS OF ORDER 4 HAVE A COMMON SQUARE.

By G. A. Miller.

If a group G involves only operators whose orders divide 4 and all of its operators of order 4 have a common square then its quotient group with respect to the subgroup generated by this square involves only operators of order 2 and hence this subgroup includes the commutator subgroup of G. It results that G is either abelian or has a commutator subgroup of order 2. In the former case G is either of type $(1, 1, 1 \cdots)$ or of type $(2, 1, 1, \cdots)$, and its order is always of the form 2^m. If the operators of order 2 contained in G do not generate G they must generate an invariant subgroup of G which appears in its central since all of the remaining operators of G are of order 4 and must generate G. If such an operator were not commutative with an operator of order 2 contained in G its product into this operator of order 2 would be of order 2. As this is impossible it results that *a necessary and sufficient condition that a non-abelian group which involves only operators whose orders divide 4 and whose operators of order 4 have a common square contains a non-invariant operator of order 2 is that it is generated by its operators of order 2.*

From this theorem it results that when G does not involve a non-invariant operator of order 2 it is either abelian or Hamiltonian. As these groups are well known we shall assume in what follows that G involves a non-invariant operator s_1 of order 2 and we shall represent by H_1 the subgroup of index 2 under G which is composed of all the operators of G which are separately commutative with s_1. If H_1 contains an operator s_2 of order 2 which is non-invariant under H_1 then we represent by H_2 the subgroup of index 2 under H_1 composed of all the operators of H_1 which are commutative with s_2. By continuing this process we obtain a set of commutative operators of order 2 $s_1, s_2, \cdots, s_\lambda$ which appear in an invariant subgroup involving only invariant operators of order 2. This subgroup must therefore belong to one of the following three types: Hamiltonian, abelian and of type $(1, 1, 1, \cdots)$, abelian and of type $(2, 1, 1, \cdots)$.

Exactly half of the operators of G which do not appear in H_1 are of

order 2, while the rest of these operators are of order 4 since the order of the product of s_1 into such an operator of order 4 is 2 and the order of the product of s_1 into such an operator of order 2 is 4. Since similar remarks apply to the invariant subgroups $H_2, \cdots, H_\lambda$ contained in the invariant subgroups $H_1, H_2, \cdots, H_{\lambda-1}$, respectively, it results that a necessary and sufficient condition that we arrive at a Hamiltonian group by the given process is that more than one-half of the operators of G are of order 4. If less than one-half of these operators are of order 4 we thus arrive at an abelian group of type $(1, 1, 1, \cdots)$, while if exactly one-half of these operators are of order 4 we arrive at an abelian group of type $(2, 1, 1, \cdots)$. That is, the selection of the separate operators $(s_1, s_2, \cdots, s_\lambda)$ does not affect the type of subgroup reached by the given process.

The subgroup generated by s_1 is not invariant under G but it is invariant under H_1. Similarly, the subgroup generated by s_1, s_2 is not invariant under H_1 but it is invariant under H_2, otherwise G would be the direct product of H_2 and the subgroup generated by s_1 and s_2. In general, the subgroup generated by $s_1, s_2, \cdots, s_\alpha$, $\alpha \leqq \lambda$, is not invariant under $H_{\alpha-1}$ but it is invariant under H_α, where $H_0 = G$. Hence H_α, $\alpha < \lambda$, is the direct product of a subgroup of $H_{\alpha+1}$ and the group generated by $s_1, s_2, \cdots, s_\alpha$. Moreover, $H_{\alpha-1}$ is obtained by extending this direct product by an operator of order 2 which is commutative with every operator of $H_{\alpha+1}$ but not with s_α. The subgroup H_α is the cross-cut of the subgroups of index 2 composed of all the operators of G which are separately commutative with $s_1, s_2, \cdots, s_\alpha$, respectively, and exactly half of the operators in each of the corresponding co-sets of G except H_α are of order 2 while the rest are of order 4 since each of these co-sets is composed of operators which are not commutative with an operator of order 2 contained in G.

From what precedes it results directly that there are three infinite categories of groups coming under the heading of the present article and that the number of the groups of order 2 in each of these categories increases with the increase of the value of m. For every value of $m > 1$ there are two such abelian groups. For every value of $m > 3$ there is also at least one non-abelian group in each of the three given categories. When H_λ is the abelian group of type $(1, 1, 1, \cdots)$ it results that $m - \lambda = \lambda + 1$ and hence $m = 2\lambda + 1$. The number of the non-abelian groups which belong to this category is obviously equal to the number of the distinct natural numbers which can be assigned to λ in this equation. In particular, when $m = 3$ the octic group is the only non-abelian group which belongs to this category while the abelian group of order 8 and of type

(1, 1, 1) is the abelian group which comes thereunder. For every value of $m > 2$ this category includes one and only one group in which exactly one-fourth of the operators are of order 4.*

When H_λ is abelian and of type $(2, 1, 1, \cdots)$ then $m - \lambda = \lambda + 2$ since this abelian group must be of an order which is at least equal to 4. Hence in this case $m = 2\lambda + 2$ and there is one and only one such non-abelian group for every natural number which satisfies this equation. When H_λ is a Hamiltonian group then $m - \lambda = \lambda + 3$ since the order of a Hamiltonian group is at least 8. Hence $m = 2\lambda + 3$ and all the groups which belong to this category are non-abelian. That is, the number of non-abelian groups in this case is one more than the number of the distinct natural numbers which can be substituted for λ in the equation $m = 2\lambda + 3$. Hence the following theorem has been established: *The number of the non-abelian groups of order 2^m involving only operators whose orders divide 4 and whose operators of order 4 have a common square is one more than the number of the possible solutions which can be obtained by substituting distinct natural numbers for λ separately in the following three equations:* $m = 2\lambda + 1$, $m = 2\lambda + 2$, $m = 2\lambda + 3$. In particular, when $m = 6$ there are 6 such non-abelian groups † that is, there are exactly 8 groups of order 64 which involve only operators whose orders divide 4 and whose operators of order 4 have a common square.

Whenever H_λ is abelian it is the direct product of the central of G and the group generated by $s_1, s_2, \cdots, s_\lambda$, when it is non-abelian its operators of order 2 generate its central. Every subgroup of G which involves operators of order 4 is invariant under G. In particular, a non-invariant subgroup of G is abelian, and of type $(1, 1, 1, \cdots)$ but such a subgroup is not always non-invariant. The given systems of groups may obviously be regarded as a generalization of the Hamiltonian groups of order 2^m. While the subgroup of index 2 composed of all the operators of such a G which are commutative with a non-invariant operator of order 2 belongs always to the same category as the original group the subgroup of index 2 composed of all the operators of G which are commutative with one of its non-invariant operators of order 4 must always belong to the category in which exactly half the operators are of order 4.

When such a group G is not generated by its operators of order 4 contained therein these operators generate an invariant subgroup of G which

---

* G. A. Miller, Paris *Comptes Rendus*, Vol. 141 (1905), p. 891.

† G. A. Miller, *American Journal of Mathematics*, Vol. 52 (1930), p. 634.

must be abelian since each of its operators is transformed into its inverse by every operator of order 2 which does not appear in this subgroup. Hence there is one and only one such G for every value of $m > 2$, and this infinite system is characterized by the fact that exactly one-fourth of its operators have an order which exceeds 2, as was noted above. This system is composed of all the groups coming under the heading of the present article which involve operators of order 4 and are separately generated by their operators of order 2 but not by their operators of order 4. Those which are generated by their operators of order 4 but not by their operators of order 2 must be either Hamiltonian or abelian and of type $(2, 1, 1, \cdots)$. Each of the other groups which comes under this heading is therefore generated both by the operators of order 2 contained therein and also by the operators of order 4 which it involves.

COVERING THEOREMS IN GENERAL TOPOLOGY.

By Selby Robinson.

1. *Introduction.* E. W. Chittenden † found two necessary and sufficient conditions that a set E of a general topological space (P, K) have the property of Borel-Lebesgue in the following form: every infinite proper covering $\mathfrak{F}$ of E is reducible.‡ The present investigation began with the discovery of properties of sets in a general space (P, K) which are equivalent to the property: every infinite *proper* covering of the entire space P is reducible to a *simple* covering of E. In this property we introduce two new concepts; reducibility of a covering of P to a covering of the set E, and the reducibility of a covering of one type to one of another. The investigation as a whole turns about the relations between a variety of reducibility properties and corresponding properties, some of which are related to boundedness and the others to the property: there is a point common to all sets of a descending sequence of closed subsets of an interval.

In sections 2 to 4 below we formulate a number of theorems of special interest which, with the exception of Theorem 1, are corollaries of the more general theorems of section 6. These theorems in turn follow readily from the results of a study of the reducibility of a family of subsets of a set E made by E. W. Chittenden and myself.§ These theorems of section 6 present equivalences involving the reducibility of coverings of a set of points H of a general topological space which are of one abstract type to coverings of a set E which are of the same or a different type. In particular, the set H may be the set E as in the property of Borel-Lebesgue or all points of the space as in the other reducibility property of the preceding paragraph.

A group of equivalences studied by Sierpinski is extended in sections 9 and 10. In conclusion it is shown in section 10 how a principle of duality introduced in "Reducibility" leads to an analogy between separability and the property of Lindelöf.

† "On general topology, etc.," *Transactions of the American Mathematical Society*, Vol. 31 (1929), p. 306. This paper is hereafter referred to as "Topology."

‡ A family $\mathfrak{F}$ of sets V is a proper covering of E if each point of E is *interior* to some set V of $\mathfrak{F}$, a *simple* covering if each point of E is an element of some set V of $\mathfrak{F}$. A proper covering of a set E is *reducible* if some subfamily $\mathfrak{F}_1$ of $\mathfrak{F}$ of lower power is also a proper covering of E. If the points of E are elements of the sets of some subfamily $\mathfrak{F}_1$ of power less than $\mathfrak{F}$, $\mathfrak{F}$ is reducible to a simple covering.

§ "On the reducibility of families of subsets and related properties," *American Journal of Mathematics*, Vol. 55 (1933). This paper is hereafter referred to as "Reducibility." In Theorem 10 of "Reducibility," read, Let $\lambda = \aleph_0$."

2. *Reducibility of proper coverings to simple ones.* Consider the following properties of the sets E and H of a general topological space P.

A. Every infinite subset of regular power E is nuclear † *in H.*

B. Every infinite decreasing sequence of subsets of E is closed in H.

C. Every infinite proper covering of $\mathfrak{F}$ of H is reducible to a simple covering of E of lower power.‡

THEOREM 1. *If H contains E,*§ *properties A, B, and C are equivalent.*¶

We shall prove that $B \rightarrow A \rightarrow C \rightarrow B$. To prove that B implies A; suppose that an infinite subset Q of E of regular power were not nuclear in H. Arrange the points of Q in a sequence q_α where α ranges over all ordinals $0 < \alpha < \Omega(/Q/)$, the least ordinal of power $/Q/$. For any such α let $G_\alpha = \sum_{\alpha' \geqq \alpha} q_{\alpha'}$. The decreasing sequence $\mathfrak{S} = [G_\alpha]$ is closed in H, or we might say closed in some point q of H. Then q is a nuclear point of Q. For if there were a neighborhood V of q which contained less than $/Q/$ points of Q, the points of $Q \cdot V$ cannot run through the sequence q_α since $/Q/$ is regular.

To prove A implies C, suppose on the contrary that an infinite proper covering $\mathfrak{F}$ of H is not reducible as required. Let the sets of $\mathfrak{F}$ be arranged in a sequence V_α, $0 < \alpha < \Omega(/\mathfrak{F}/)$. Let $E_\alpha = V_\alpha - \sum_{\alpha' < \alpha} V_{\alpha'}$. At least $/\mathfrak{F}/$ of the sets E_α are non-null. It is therefore possible to choose a series of indices $0 < \beta < \Omega(\nu)$ where ν is *regular* such that for every α there is an index α_β such that $\alpha < \alpha_\beta$, and a set Q of distinct points q_β of E such that if $\alpha < \alpha_\beta$ then q_β is not an element of V_α. By hypothesis some point q of H is a nuclear point of Q, that is every neighborhood of Q contains ν points of Q, but by construction no V_α can contain ν points of Q.

Suppose C holds but there is a decreasing sequence $\mathfrak{S}$ of subsets of E which is not closed in H. Then there is a subsequence ‖ $\mathfrak{S}_\mu = [G_\alpha]$ of regu-

† Our definition of nuclearity of a set is different from that of "Topology." A set A is *nuclear* in a set H if there is a point q of H such that every neighborhood of q contains $/A/$ points of A. A decreasing sequence $\mathfrak{S}_\mu = [G_\alpha/0 < \alpha < \Omega(\mu)]$ of subsets of E is *closed* in H if every neighborhood of some point q of H *contains* a point of each set G_α of $\mathfrak{S}_\mu$.

‡ We have an example which shows that this property is weaker than the property; every proper covering of H is reducible to a finite simple covering.

§ The hypothesis, H contains E can be omitted if property C is generalized by requiring only that $\mathfrak{F}$ be reducible to a simple covering of $E - D$ of power less than $/\mathfrak{F}/$, where D is a subset of E (perhaps null) of power $< /\mathfrak{F}/$.

¶ The equivalence of A and B is stated on page 307 of "Topology," for the case $H = P$.

‖ The notation $\mathfrak{S}_\mu$ will always designate a decreasing sequence of subsets of E, which is order type $\Omega(\mu)$, where μ is some infinite cardinal not necessarily regular.

lar power which is not closed in H. For any α less than $\Omega(\mu)$, let $V_\alpha = P - G_\alpha$. Let $\mathfrak{F} = [V_\alpha]$. For any point h of H there is an index α such that h is not in $G_\alpha + L(G_\alpha)$. Hence h is interior to V_α. Then since $\mathfrak{F}$ is a proper covering of H, there is a subfamily $\mathfrak{F}_1 = [V_{\alpha_\beta}]$ of $\mathfrak{F}$ of regular power less than $/\mathfrak{F}/$ which covers E simply. Since μ is regular and $/\mathfrak{F}_1/$ less than μ, there is an index $\gamma < \Omega(\mu)$ which is greater than any α_β. Then no point of G_γ is contained in any set of $\mathfrak{F}_1$, which thus cannot cover E simply.

By the same reasoning, a similar theorem can be established for enumerably infinite coverings.

THEOREM 2. *If H contains E, the following properties are equivalent:*

A. Every enumerable infinite subset of E has a nuclear point in H.

B. Every enumerable infinite decreasing sequence of subsets of E is closed in H.†

C. Every enumerable proper covering of H is reducible to a finite simple covering of E.

This theorem holds if some other regular cardinal μ be substituted for $\aleph_0$. Consider the case $\mu = \aleph_1$ and for simplicity let $H = E$. Then the A property takes the form, E is self-condensed.‡ Since this property is equivalent to the reducibility of proper coverings of power $\aleph_1$ to enumerable simple ones, it is implied by the Lindelöf property, every proper covering of E is reducible to an enumerable one. Fréchet asks for conditions under which the two properties are equivalent. They evidently are equivalent if the interior of every set in the space is open and $/E/ = \aleph_1$; § or more generally if the interior of every set is open and every proper covering of E of power greater than $\aleph_1$ is reducible.

For μ irregular we can in general say only that A implies B which implies C. But if $\mu = /H/$, we have proved that properties A and B are equivalent. For $/H/$ irregular, the proof is a generalization of one given by Sierpinski.¶ The same equivalence extends to nuclearity and closure in

† For $H = P$ or $H = E$ and P a neighborhood space, the equivalence of A and B was proved by Fréchet, *Les Espaces Abstraits*, Paris (1928), p. 231; and *American Journal of Mathematics*, Vol. 50 (1928), p. 52.

‡ Every non-enumerable subset Q of E has a point of condensation q in E; i.e., every neighborhood of q contains a non-enumerable number of points of Q. Fréchet, *Espaces Abstraits*, p. 174.

§ Contrary to a statement of Fréchet, *Espaces Abstraits*, p. 234.

¶ *Bulletin of the American Mathematical Society*, Vol. 32, p. 652. Cf. Chittenden, *Bulletin of the American Mathematical Society*, Vol. 30, p. 514. Chittenden and Sierpinski proved that if every infinite decreasing sequence of subsets of E is closed in E, then E has a nuclear point in E. But Sierpinski showed by an example that

coverings. In that case the requirement that H shall be of the same power as the subsets of E which are to be proved nuclear is replaced by the requirement that the coverings shall be of that power. Many theorems of this paper are based on this idea. It is used in the proof of Theorem 2 of "Reducibility" and Theorem 3 of this paper may be regarded as a corollary of that theorem.

THEOREM 3. *A necessary and sufficient condition that every subset of E of power $/H/$ have a nuclear point in H, is that every sequence $\mathfrak{S}_{/H/}$ of subsets of E shall be closed in H.*

It is interesting to compare the properties of Theorem 1 with the following property $C4$.† *Every proper covering of H is reducible to a finite simple covering of E.* We make use of the concept of the nuclearity of a set Q in a family $\mathfrak{F}$ of sets V. This means that some set V contains $/Q/$ points of Q. Likewise, a sequence $\mathfrak{S}$ of sets G is closed in $\mathfrak{F}$ if there is a single set V of $\mathfrak{F}$ which contains a point of each set G.‡ The following theorem is a consequence of Theorem 11 of section 6.

THEOREM 4. *If H contains E, property $C4$ implies the properties of Theorem 1 and is equivalent of the following properties:*

A. Any infinite proper covering of H has an enumerable subfamily in which every enumerably infinite subset of E is nuclear.

B. And infinite proper covering of H has an enumerable subfamily in which every enumerable decreasing family of subsets of E is closed.

3. *Reducibility of proper coverings to proper coverings.* The following theorem should be compared with Theorem 1 and the next with Theorem 2. Both are corollaries of Theorem 10 of section 6.

THEOREM 5. *If H contains E, the following properties are equivalent:*

A. Every infinite regular subset of E of regular power has a proper nuclear point in H.

B. Every infinite decreasing sequence of subsets of E, is properly closed in H.

C^*.§ Any infinite family $\mathfrak{W}^*$ of sets such that each point h of H is

---

every infinite decreasing sequence of subsets of E might be closed in the space containing E but E have no nuclear point in the space.

† Property C of Theorem 4.

‡ "Reducibility," section 2.

§ A family $\mathfrak{W}^*$ can also be regarded as the sum of the interiors of the sets of a proper covering of H. See Lemma 3, section 6.

interior to some set V_h whose interior is contained in a set of $\mathfrak{W}^*$, is reducible to a simple covering of E of power less than $/\mathfrak{W}^*/$.

THEOREM 6. *If $H > E$, the following properties are equivalent:*

A. Every enumerably infinite subset of E has a proper nuclear point in H.

B. Every enumerable decreasing sequence of subsets of E is properly closed in H.

C. Any enumerable family $\mathfrak{W}^*$ of sets such that each point h of H is interior to some set V_h whose interior is contained in a set of W^*, is reducible to a finite simple covering.*

Fréchet proposed the problem of finding a nuclearity property equivalent to the Borel property, after Sierpinski had shown that the Borel property does not imply that every enumerable subset of E has a nuclear point in E.† A solution of this problem is presented in the following theorem.

THEOREM 7. *If $H > E$, the following properties are equivalent:*

A. Every enumerable subset of E is property nuclear in every enumerable proper covering of H.

B. Every enumerable decreasing sequence of subsets of E is properly closed in every enumerable proper covering of H.

C. Every enumerably infinite proper covering of H is reducible to a finite proper covering of E.‡

When $\mu = /H/$, we secure a result which like Theorem 3 depends upon our generalization of Sierpinski's proof.

THEOREM 8. *A necessary and sufficient condition that every subset of E of power $/H/$ have a proper nuclear point in H, is that every sequence $\mathfrak{S}_{/H/}$ of subsets of E be properly closed in H.*

Theorem 7 is a consequence of Theorem 12. It is also possible to secure from Theorem 12 necessary and sufficient conditions that every proper covering of H be reducible to a proper covering of E of lower power. Necessary and sufficient conditions (similar to those of Theorem 4) can be found for the reducibility of all proper coverings of H to *finite* proper coverings of E.

4. *Reducibility of proper coverings of E.* There are relations among the properties of the last section, which hold only when $H = E$. These additional relations are consequences of the fact that reducibility of proper coverings of E of power greater than or equal to μ is equivalent to reducibility to

† *Les Espaces Abstraits*, pp. 230-231.

‡ In an $\mathfrak{L}$-space having the property of Borel, the interior of every set is open so that property $A2 = A6$. Hence the properties of Theorem 7 are equivalent to those of Theorem 6 in an $\mathfrak{L}$-space.

power less than μ. On page 306 of "Topology," Chittenden states that property $B5$ is equivalent to the reducibility of every infinite proper covering of E, and to nuclearity in E of every infinite subset of E.† We also obtain the following result which is invalid when $H \neq E$.

THEOREM 9. *The following properties are equivalent:*

A. Every subset of E of power $\geqq \lambda$ has a proper nuclear point in E.

B. Every sequence $\mathfrak{S}_\mu$, $\mu \geqq \lambda$, is properly closed in E.‡

That A implies B can be proved by the method of Theorem 1. The proof that B implies A is related to the proof of Theorems 3 and 8, and makes use of the theorems on proper nuclearity and closure derived from lemmas 1 and 2 of section 6. We wish to show that if B holds the subset Q of E of power greater than or equal to λ is properly nuclear in any proper covering $\mathfrak{F}$ of E. Property B implies that $\mathfrak{F}$ is reducible to a proper covering $\mathfrak{F}_1$ of power less than λ. Now Sierpinski's proof can be extended to show that for any infinite cardinal μ, proper closure of all sequences $\mathfrak{S}_\mu$ in a family $\mathfrak{F}_1$ of power $\leqq \mu$, implies proper nuclearity in $\mathfrak{F}_1$ of all subsets of power μ. Then the set Q in question is properly nuclear in the subfamily $\mathfrak{F}_1$ of $\mathfrak{F}$ above mentioned, hence properly nuclear in $\mathfrak{F}$.

In the same way, proper closure of every sequence $\mathfrak{S}_\mu$, $\lambda \leqq \mu \leqq \nu$, of subsets of E in every proper covering of E whose cardinal number is in that interval, is equivalent to proper nuclearity of subsets of E of a power which is in the given interval in proper coverings of any power in the interval.

5. *The set functions I and T.* The theorems of the preceding section can be generalized by the consideration of the reducibility of coverings of H of an abstract type I to coverings of E of type T. An example already considered is that of the reducibility of proper coverings to simple coverings. A point q is an I_T-point of a set G if every set which is in the relation I to q is in the relation T to a point of G. Corresponding to reducibility of I-coverings to T-coverings, we may consider I_T-closure in H of sequences $\mathfrak{S} = [G]$, which is defined as the property that there is in H a point q which is an I_T-point of each G. For any set G, denote by $I_T(G)$ the set of all I_T-points

† Compare property $A5$.

‡ For a Hausdorff space with $H = E$, Hildebrandt incorrectly states (*Bulletin of the American Mathematical Society*, Vol. 32 (1926), p. 468), that property A is equivalent to the reducibility of all proper coverings of E to power less than λ and to a property B somewhat similar to $B9$. As a matter of fact Hildebrandt's properties A and B are each stronger in Hausdorff spaces than his reducibility property. Because his B property is stated in terms of well ordered decreasing sequences instead of sequences $\mathfrak{S}_\mu$, it is satisfied in a Hausdorff space which does not have property A.

of G. The set function $I_T(G)$ determined by the set functions I and T is monotonic. That is, if $G_2 > G_1$ then $I_T(G_2) > I_T(G_1)$.

For the special case $T(G) = G$, we designate $I_T(G)$ by the symbol $I_0(G)$. When the set function $I(G)$ is monotonic, $I_0(G) = CIC(G)$, so that $I(G) = CI_0C(G) = (I_0)_0(G)$. Thus there is symmetry between $I(G)$ and $I_0(G)$. If $I(G)$ is the set of all points interior to G in the ordinary sense, $I_0(G) = M(G)$† $= G + L(G)$.‡

The point q is an I_T-nuclear point of the set G, if every I-set of q is in the relation T to $/G/$ points of G.

6. *A general theory of reducibility.* A set Q may be said to be T-nuclear in a family $\mathfrak{F}$ of sets if there is a set of $\mathfrak{F}$ which is in the relation T to $/Q/$ points of Q. A sequence, or more generally any family $\mathfrak{S}$, of sets is T-closed in a family $\mathfrak{F}$ if there is some set of $\mathfrak{F}$ which is in the relation T to a point of each set of $\mathfrak{S}$. These properties are related to nuclearity and closure in a set in the following manner.

LEMMA 1. *A necessary and sufficient condition that an infinite set Q have an I_T-nuclear point in the set H, is that Q be T-nuclear in every I-covering of H.*

LEMMA 2. *A necessary and sufficient condition that the family $\mathfrak{S}$ of sets be I_T-closed in H, is that $\mathfrak{S}$ be T-closed in every I-covering of H.*

A somewhat similar equivalence holds for the C^* properties.

LEMMA 3. *A necessary and sufficient condition that a family $\mathfrak{W}^*$ of sets be such that every point h of H shall have an I-neighborhood whose T-set is contained in some set of $\mathfrak{W}^*$, is that there be an I-covering $\mathfrak{F} = [V]$ of H such that every set $T(V)$ is contained in some set of $\mathfrak{W}^*$.*

Evidently the condition in lemma 1 is necessary. For if the point q of H is an I_T-nuclear point of Q, in every I-covering of H there will be an I-neighborhood of q which will therefore contain $/Q/$ points of Q. Suppose that the converse is not true. Then there is for each point h of H an I-neighborhood V_h of h which is not in the relation T to $/Q/$ points of Q. Then the family of sets V_h is an I-covering of H in which Q is not T-nuclear. The other lemmas are proved in a similar manner.

Consider a family $\mathfrak{F}$ of subsets V of a general topological space P, a subset E of P, a relation T between subsets and points of P, and an infinite

† "Topology," p. 295.

‡ If from a given pair of set functions I and T, we form I_T and from it form $(I_T)_0 = CI_TC$, the latter function occurs in the C^* properties. For the hypothesis on a family $\mathfrak{W}^*$ of sets W^* is that each point H is in some set $CI_TC(W^*)$.

cardinal μ. In the study of T-nuclearity of subsets of E and T-closure of sequences $\mathfrak{S}$ in this family $\mathfrak{F}$, in relation to the reducibility of $\mathfrak{F}$ to power less than μ; the family $\mathfrak{F}$ can be replaced by the family $\mathfrak{W}$ of sets $W = E \cdot T(V)$. Thus, in the presentation in "Reducibility," the relation T, the space P, and the family $\mathfrak{F}$, did not appear, but only the cardinal μ and the family $\mathfrak{W}$ of subsets of E. The theorems obtained concerning μ and $\mathfrak{W}$ were next extended to an arbitrary class $\boldsymbol{W}$ of families $\mathfrak{W}$ and an arbitrary class M of infinite cardinals. It is possible to use the relation T in the statement of these theorems, substituting for $\boldsymbol{W}$ the class $\boldsymbol{F}$ of families $\mathfrak{F}$. From theorems 4, 7, 8, and 9 of "Reducibility" for the special case in which $\boldsymbol{F}$ is the class of all I-coverings of H, and from lemmas 1, 2, and 3; we derive the following theorems:

THEOREM 10. *Each of the following properties implies its successor: If all cardinals in M are regular, all the properties are equivalent.*

A. Every subset of E whose power is in M, has an I_T-nuclear point in H.

B. Every sequence $\mathfrak{S}_\mu$, μ in M, is I_T-closed in H.

C^*. Any family $\mathfrak{W}^*$, $/\mathfrak{W}^*/$ in M, which has the property that any point h of H has an I-neighborhood whose T-set is contained in a set of $\mathfrak{W}^*$, has a subfamily of lower power which covers E simply, except for a subset of power less than $/\mathfrak{W}^*/$.†

COROLLARY. *If M is the class of all infinite cardinals less than ν, properties $B10$ and C^*10 are equivalent to the property: any subset E of infinite regular power less than ν has an I_T-nuclear point in H.*

Clearly Theorem 5 is a consequence of this corollary.

THEOREM 11. *The following properties are equivalent:*

A. Any I-covering of H contains a subfamily $\mathfrak{F}'$ of power $\leqq \mu$ in which every subset of E of power $/\mathfrak{F}'/$ is T-nuclear.

B. Any I-covering of H contains a subfamily $\mathfrak{F}'$ of power $\leqq \mu$ in which every sequence $\mathfrak{S}_{/\mathfrak{F}'/}$ is T-closed.

C. Any I-covering of H has a subfamily of power $< \mu$ which is a T-covering of E except for a set of power less than μ.

THEOREM 12. *The following properties are equivalent:*

† We were able to omit this last phrase from the statements of the C^* and C properties in previous theorems because the I-relation was always at least as strong as the T-relation and because we made the hypothesis, H contains E. Then the subfamily which covers all but $/\mathfrak{W}^*/$ points of E can be enlarged by the addition of sets of $\mathfrak{W}^*$ which cover the other points of E, so that we secure a family which is still of power less than $/\mathfrak{W}^*/$ but which covers E.

A. *Any I-covering of H whose power is in M, has a subfamily $\mathfrak{F}'$ in which every subset of E of power $/\mathfrak{F}'/$ is T-nuclear.*

B. *Any I-covering of H whose power is in M, has a subfamily $\mathfrak{F}'$ in which every sequence $\mathfrak{S}_{/\mathfrak{F}'/}$ is T-nuclear.*

C. *Any I-covering of H whose power μ is in M has a subfamily of lower power which is a T-covering of E except for a set of power less than μ.*

The theory of the reducibility of proper coverings of H to proper coverings of E, presented in section 3, is the same as the general theory of the reducibility of I-coverings of H to T-coverings of E.† The special relations discussed in section 4 hold whenever $H = E$ and I is the same as T or weaker than T. The situation described in section 2 differs from that of sections 3 and 4 in that the additional relation $C^*10 = C12$ is present, and $B10$ is equivalent to the property; every sequence $\mathfrak{S}_\mu$, μ in M, of subsets of E is closed in every proper covering of H of power μ.‡ These equivalences hold whenever;

1. $T(V) = V$ for every set V; and
2. $I(\Sigma V)$ contains $\Sigma I(V)$, for every collection of sets V.§

7. *Other choices of I and T.* Alexandroff studied¶ the reducibility of coverings of E consisting of open sets. Since the relations I and T are the same, theorems corresponding to those of section 4 hold when $H = E$. Considering more generally the reducibility of families of open sets which cover H to coverings of E of the same kind, it is easily seen that the conditions for the equivalence of C^*10 and $C12$ are fulfilled. So relations like those of section 2 are also present. In a space in which the interior of every set is open, the properties involving open sets are, respectively, equivalent to those of sections 2 and also to those of sections 3 and 4.

But Alexandroff and Urysohn introduce a kind of reducibility which is not equivalent to the preceding kinds in Hausdorff spaces. When extended to general topological spaces, this kind of reducibility splits into four kinds, determined by the combinations which occur when an I-covering is either a proper covering or a covering of open sets and $T(G)$ is either $M(G)$ or

---

† If the C^* and C properties are amended as indicated in the preceding footnote.

‡ This property is equivalent to the property any infinite subset Q of E, $/Q/$ in M, is nuclear in every proper covering of H of power $/Q/$.

§ The two conditions can be expressed in the following slightly more general form. 1. If $V' < \Sigma V$, $T(V') < \Sigma T(V)$; 2. For any ΣV, there is a V' contained in ΣV, such that $I(V') > \Sigma I(V)$.

¶ See Fréchet, *Espaces Abstraits*, p. 225 and p. 230.

$M^0(G)$.† In none of the four situations do we have C^*10 equal to $C12$ or I the same as or weaker than T. So the only relations which hold are those of section 3.

It might be of interest to consider the more general definition of neighborhood in which V may be a neighborhood of p although p is not a point of V. We remarked that if $I(G)$ means the points interior to G in the ordinary sense, $I_0(G) = CIC(G) = M(G) = G + L(G)$. If $I(G)$ means the points of which G is a neighborhood in this more general sense, $I_0(G) = L(G)$. The reducibility of coverings of neighborhoods in this more general sense to simple coverings, is related to the corresponding nuclearity and closure properties in the manner set forth in section 2. The theory of the reducibility of these more general neighborhood coverings to coverings of the same type, takes the form indicated in sections 3 and 4; and the same thing is true of the reducibility of such coverings to ordinary proper coverings. But reducibility of proper coverings to the general neighborhood coverings would seem to have only the relationships corresponding to those of section 3.

8. *Definitions of the closure of a sequence.* Fréchet gave a definition of the closure of a decreasing sequence $\mathfrak{S}$ of sets which suggests the consideration of the property; there is a point of H common either to the sets G of $\mathfrak{S}$ or to their L-sets. This is evidently equivalent to the property; there is a point of H common to the sets $G + L(G)$. With regard to the closure of all decreasing sequences of subsets of E which are of a given type, this definition is equivalent to a simpler one.

THEOREM 13. *A necessary and sufficient condition that for every decreasing sequence $\mathfrak{S}$ of order type τ* ‡ *of subsets G of E, there be a point of H common to the sets $L(G)$; is that for every sequence of that sort there be a point of H common to the sets $G + L(G)$.*

It is obvious that the first condition implies the second. To prove the converse, suppose there is a sequence $\mathfrak{B}$ of order type τ of subsets B of E such that the sets $L(B)$ have no point of H in common. Corresponding to each set B, let $G = B - \Pi B$. Let $\mathfrak{S} = [G]$. Then there is no point of H common to the sets $G + L(G)$, contrary to hypothesis.

The set $V(G)$ may be defined as the set of points p such that $L(G - p)$ contains p.§ Then if there is a point of H common to the sets $V(G)$ there

† Where $M(G) = G + L(G)$ and $M^0(G)$ is the least L-closed set containing G. "Topology," p. 295. The reducibility of proper coverings to M^0-coverings, is equivalent to reducibility of families of L-closed sets to which the points of H are interior to simple coverings of E.

‡ The order type τ is assumed to be such that there is no last set of $\mathfrak{S}$.

§ "Topology," p. 296.

must be one common to all the sets $L(G)$. Conversely, if for each sequence $\mathfrak{S}$ of order type τ of subjects G of E, there is a point of H common to the sets $L(G)$; there is for each $\mathfrak{S}$ a point of H common to the sets $V(G)$. For suppose there is a sequence $\mathfrak{B} = [B]$ of type τ such that the sets $\Pi V(B)$ have no point of H in common. Then $H \cdot \Pi V(G) = 0$, where for each set B, $G = B - \Pi B$. Then the sets $L(G)$ have no point of H in common, contrary to hypothesis. The same sort of reasoning can be used to prove that there is a point of H common to the L-sets of the sets of any sequence of subsets of E of type τ; under the hypothesis that the property holds for such sequences of that type as have no points in common.

The definition of closure, there is a point of H common to the sets $G + L(G)$, is equivalent to the definition which we gave in an earlier section. For $G + L(G)$ is the set of all points p such that every set to which p is interior contains a point of G. If $G + L(G)$ be denoted by $M(G)$, the set $M_0(G) = CMC(G) = C(C(G) + LC(G)) = C(C(G) + CL_0(G)) = G \cdot L_0(G)$. Now $L_0(G)$ is the set of all points p such that G is a neighborhood of p in the general sense in which the neighborhood is not required to contain p, and the product of this set and G is the set of points interior to G in the ordinary sense. Since interiority in the usual sense corresponds to the function $M(G)$ while the use of the general neighborhood not necessarily containinng the points of which it is a neighborhood corresponds to the function $L(G)$, by Theorem 13 I_T-closure when T means contains and I means interior is equivalent to I_T-closure with T contains and I the more general neighborhood relation. If μ is regular, it follows from this that reducibility of proper coverings of H of power μ to simple coverings of E (except for a subset of power μ), is equivalent to reducibility of the more general type of neighborhood coverings to simple coverings. Therefore the two sorts of reducibility are equivalent if both are stated for all cardinals for all less than ν. We have not been able to establish the equivalence for irregular cardinals. The two types of nuclearity are obviously equivalent.

Theorem 13 and the equivalences just following it hold for any of the types of closure defined by two monotonic set functions I and T. For the set function I_T is monotonic and may be identified with the L of Theorem 13. Then the corresponding $V(G)$ may be regarded as the set of points p such that every I-neighborhood of p is in the relation T to a point of G other than p. That closure of all sequences $\mathfrak{S}$ of order type τ in this sense is equivalent to closure in the original I_T sense, is therefore a consequence of the equivalence of the L and V types of closure, and may also be established directly by the same argument.

The next theorem relates to closure of sequences $\mathfrak{S}_\mu$ for which each set G_a of $\mathfrak{S}_\mu$ has a point which is not in $L(G_{a+1})$. The same arguments as before show that for sets of all sequences of this sort to have a point of H common to their L-sets is equivalent to their having a point of H common to their M-sets or common to their V-sets. By taking the set function L to be M, we can see that the theorem holds when the sequences considered are those in which the set G_a has a point not in $M(G_{a+1})$ rather than not in $L(G_{a+1})$.

THEOREM 14. *If μ is regular and H contains E, the closure in H of all sequences $\mathfrak{S}_\mu$ is implied by the closure of those for which each set G_a has a point which is not in $M(G_{a+1})$.*

We have an example of a set $E = H = P$ which shows that the equivalence of Theorem 14 does not hold for μ irregular.† Indeed, this set E does not have property III or property II of Theorem 16 (where I means interior to and T means contains). So the same example shows that the properties of Theorem 16 are weaker than those of Theorem 15.

The proof of Theorem 14 is made by assuming a sequence of $\mathfrak{S}_\mu = [G_a]$ not closed in H. Then for any point q of any $(G_a - G_{a+1})$ there is a set of higher index which does not have q for an L-point, since otherwise q would be in every $L(G_a)$. In this way a subsequence is secured which has the desired property and is not closed in H. The closure of the sequences having this property implies property C, even when μ is irregular.

9. *The validity of the properties for all subsets of E.* Sierpinski has studied the property; every decreasing sequence of closed sets is enumerable. This suggests a generalization in which it is asserted that there exists no sequence $\mathfrak{S}_\mu$ of subsets of E of the special kind referred to in Theorem 14. If for every subset E_0 of E every sequence $\mathfrak{S}_\mu$ of this kind consisting of subsets of E_0 is required to be closed in E_0, then no sequence of the kind can exist. In terms of the relations I and T, no sequence $\mathfrak{S}_\mu$ of subsets of E

† The theorems of this section hold for any monotonic set function $L(G)$, in particular for $M^0(G)$. Thus from the first property of Theorem 14 we derive a closure property which is one of the open set properties discussed in the first paragraph of section 7; and from the second property of Theorem 14 the property (for $E = H = P$), there is a point common to the sets of any sequence $\mathfrak{S}_\mu$ of L-closed sets. The example referred to above shows that these properties need not be equivalent for μ irregular. For in this example the interior of every set is open, and in spaces with that property the two properties of the preceding sentence are respectively equivalent to the properties of Theorem 14. We have not determined whether or not the properties of Theorem 14 are equivalent for Hausdorff spaces. Likewise, we could not determine whether or not Hildebrandt's B property discussed in the footnote following Theorem 9, was stronger than $B9$ or independent of it.

exists such that in each set G_α of $\mathfrak{S}_\mu$ there is a point q_α which is not in $G_{\alpha+1}$ or in $I_T(G_{\alpha+1})$.† Suppose such a sequence G_α existed. For each α less than $\Omega(\mu)$ let $B_\alpha = G_\alpha - \Pi I_T(G_\alpha)$. By hypothesis no B_α is null and every one is distinct. A set B_α contains no points common to the sets $I_T(B_\alpha)$ since it contains none common to the sets $I_T(G_\alpha)$. By taking $E_0 = B_1$, we see that the sequence $\mathfrak{B}_\mu = [B_\alpha]$ contradicts the hypothesis of closure of sequences of this type in every subset E_0 of E.‡

Property *III* implies the non-existence of sequences $\mathfrak{S}_{\mu'}$ of this type, $\mu' \geqq \mu$, since the first μ sets of the sequence $\mathfrak{S}_{\mu'}$ constitute a sequence $\mathfrak{S}_\mu$. By Theorem 14, property *III* is equivalent if μ is regular to I_T-closure in every subset E_0 of E of *every* sequence $\mathfrak{S}_\mu$ of subsets of E_0. Then for any μ, property *III* implies, closure in E_0 of every sequence $\mathfrak{S}_{\mu'}$, $\mu' \geqq \mu$ and regular.

A well known property of a set E is that every non-enumerable subset A of E have a point of condensation in A. A generalization of this property is given in property *I* of the next theorem.

THEOREM 15. *The following properties are equivalent:*

B. For any subset E_0 of E and any sequence $\mathfrak{S}_\mu$ of subsets of E_0, $\mathfrak{S}_\mu$ is I_T-closed in E_0.

I. Any subset A of E of power μ, has an I_T-nuclear point in A.§

II′. Any subset Q of E of power $\geqq \mu$ has a subset X of power μ such that every point of X is an I_T-nuclear point of X.

The proof that *I* implies *II′* follows the lines laid down by Sierpinski. Let Q be any subset of E of power greater than or equal to μ, and consider any subset Y of Q of power μ. Then the desired set X is the set of all points of Y which are I_T-nuclear points of Y. The set $Y - X$ is of power less than μ, since otherwise it would by property *I* contain an I_T-nuclear point of itself, hence of Y. Therefore any point of X is an I_T-nuclear point of X.

Since property *II′* obviously implies *I*, we proceed to prove that *I* is implied by *B*. If property *B* were stated for all subsets A of E of power μ rather than for all subsets E_0, the properties would be equivalent by the

† This is referred to hereafter as property *III*.

‡ Since the set ΠG_α may be subtracted before this argument is made, the point q_α may be required not to be in every G_α. If for every subset E_0 of E, *every* sequence $\mathfrak{S}_\mu$ of subsets of E_0 is I_T-closed in E_0, then every set G_α of $\mathfrak{S}_\mu$ has a point q_α which is in $I_T(G_\alpha)$ but not in ΠG_α.

§ Properties *I, II, III* are generalizations of the corresponding properties proved equivalent by Sierpinski for spaces $\mathfrak{S}$; *Fundamenta Mathematica*, Vol. 2, p. 179; cf. Putnam, *Bulletin of the American Mathematical Society*, Vol. 36 (1930), pp. 653-654; and Robinson, *Bulletin of the American Mathematical Society*, Vol. 37 (1931), p. 629.

theorem of which Theorems 3 and 8 are special cases. However, property B as applied only to subsets A of power μ is implied by property B as stated in the theorem, being a special case. From Theorem 10, B is implied by the property; for any subset E_0 of E and any subset A of E_0 of power μ, A has an I_T-nuclear point in E_0. But the latter property is obviously implied by I.

By the same argument it follows that for every subset A of E of power greater than or equal to μ to have an I_T-nuclear point in A, it is necessary and sufficient that any sequence $\mathfrak{S}_{\mu'}$, $\mu' \geqq \mu$, of subsets of any set E_0 of E be I_T-closed in E_0. These properties are probably stronger than those of Theorem 15. These nuclearity and closure properties stated for all regular cardinals greater than or equal to μ are equivalent, but probably are weaker than the properties of Theorem 15. That these properties are implied by $B15$ can be seen from the fact that property $B15$ implies property *III*, which implies the non-existence of the sequences of $\mathfrak{S}_{\mu'}$, $\mu' \geqq \mu$, of the sort described in *III*, which implies I_T-closure of *all* sequences $\mathfrak{S}_{\mu'}$ of regular power greater than or equal to μ.

In the following theorem we consider the reducibility properties as applied to every subset of E.

THEOREM 16. *The following properties are implied by those of Theorem* 15, *and each of them implies the ones which follow. If μ is regular, all these properties are equivalent and equivalent to those of Theorem* 15.

II. Any subset of E of power $\geqq \mu$ has a subset X such that each point x of X is in $I_T(X - x)$.

III. There is no sequence $\mathfrak{S}_\mu$ of subsets of E such that each set G_α of $\mathfrak{S}_\mu$ has a point not in $G_{\alpha+1} + I_T(G_{\alpha+1})$.

C. Any I-covering of any subset E_0 of E, has a subfamily of power less than μ which is a T-covering of E_0.

C′. Every I-covering of a subset Q of E of power μ is reducible to a T-covering of power less than μ.†

It is clear that property *II′* implies property *II*. It is also easy to see that C' implies property I when μ is regular. For I_T-nuclearity in points of a subset Q of power μ is equivalent to T-nuclearity in I-coverings of Q of power

† Property C implies that every family $\mathfrak{W}^*$ of power μ related to any subset E_0 of E in the manner described in property C^*10, is reducible to a simple covering of lower power. This implies that every I-covering of power μ of any subset E_0 is reducible to a T-covering of lower power; which in turn implies property C'. These are the properties C^*10 and $C12$ applied to all subsets of E, while property C can be secured in the same way from $C11$ and property C' from reducibility of any I-covering of H to T-coverings of power less than μ of all subsets of E of power μ.

μ, which is equivalent when μ is regular to reducibility of such coverings to T-coverings of lower power.

Suppose *II* holds but not *III*. Let $\mathfrak{S}_\mu = [G_\alpha]$ be a sequence of the type prohibited by property *III*. Consider the set $Q = [q_\alpha]$ such that q_α is not in $I_T(G_{\alpha+1_1})$ or in $G_{\alpha+1}$. Let q_β be the first point of the subset X of Q whose existence is asserted by property *II*. But $(X - q_\beta)$ is contained in $G_{\beta+1}$. Since q_β is not an I_T-point of the latter set, it cannot be an I_T-point of the former; contrary to the assumption regarding X. Next suppose property *III* but not C is valid. There will be an *I-covering* $F = [V]$ of some subset E_0 of E which is not reducible to a T-covering of power less than μ. Choose a point q_1 of E_0 and a set V_1 of $\mathfrak{F}$ which is in the relation I to q_1. Then choose a point q_2 of E_0 not in the relation T to V_1 and a set V_2 of F in the relation I to q_2. Proceeding in this way we secure a set $Q = [q_\alpha]$ of order type $\Omega(\mu)$ and a corresponding sequence V_α, such that no point q_α is in the relation T to a set V_β, $\beta < \alpha$. Let $G_\alpha = \sum_{\alpha' \leqq \alpha} q_{\alpha'}$ and let $\mathfrak{S}_\mu = [G_\alpha]$. The $\mathfrak{S}_\mu$ is a sequence prohibited by *III*, since any q_α has an I-neighborhood V_α which is not in the relation T to any point of $G_{\alpha+1}$. Obviously C implies C'. It is clear that *II* and *II'* would be unchanged if stated for subsets of E of power μ rather than for all subsets of power $\geqq \mu$. Property *II* is unchanged if the sets X are required to be of power μ.

In the case in which I means interior to and T means contains and $\mu = \aleph_1$, certain of the preceding results were secured by Sierpinski in $\mathfrak{S}$ spaces; † and for the same choice of μ and I but with T also meaning interior to, by Kuratowski and Sierpinski in $\mathfrak{L}$-spaces.‡ It would be very easy to write out the results of this and the two preceding sections for any pair of relations I and T discussed in section 7.

10. *Separability.* In the article just cited Sierpinski has shown that if every set in a given $\mathfrak{S}$-space is separable, then every ascending sequence of closed sets in the space is enumerable. We have found extensions of this theorem to spaces $(P; K)$ analogous to the theorems of the preceding section. It is of interest to consider the condition for the separability of a single set E which does not require the separability of all subsets of E.

THEOREM 17. *A necessary and sufficient condition that a set E of power $\aleph_1$ be separable, is that there exist no ascending family* $\mathfrak{S}_\mu = [G_\alpha/0 < \alpha < \Omega(\aleph_1)]$

† An $\mathfrak{S}$-space is an $\mathfrak{L}$-space in which derived sets are closed. *Espaces Abstraits*, p. 211.

‡ Sierpinski, *Fundamenta Mathematica*, Vol. 2, pp. 179-188; and Kuratowski and Sierpinski, *Fundamenta Mathematica*, Vol. 2, pp. 176-178. Our proof that *III* implies C was suggested by their proof that *I* implies C.

of subsets G_a of E such that each point of E is in some G_a and every set G_{a+1} contains a point q_a not in $M(G_a)$. If $/E/ > \aleph_1$ the condition is necessary.

Proof. Suppose there is a sequence $\mathfrak{S}_\mu = [G_a]$ contradicting Theorem 17. Then E cannot be separable. For suppose some enumerable subset N of E is dense in E. There will be an index α so large that G_a contains N. Then q_a is not in $N + L(N)$, hence E is not separable. Suppose $/E/ = \aleph_1$ and E is not separable. Arrange the points of E in a sequence e_a. Rearrange these points in a sequence q_a as follows. Let $q_1 = e_1$. Let q_a be a point not in $G_a = \sum_{a' \geqq a} q_{a'}$ and not in $L(G_a)$. Then $\mathfrak{S} = [G_a]$ is a sequence contradicting Theorum 17.

That separability is closely related to the reducibility property of Lindelöf is shown by the following theorem which can be easily deduced from Theorem 11 of "Reducibility."

THEOREM 18. *The following properties are equivalent, and are present if E is separable. If $/E/ = \aleph_1$, they are equivalent to separability.*

A. For any non-enumerable family of sets to each of which a point of E is interior, there is a point q of E which is contained in $\aleph_1$ sets of the family.

B. For any decreasing sequence $\mathbf{F}_\mu$ of order type $\Omega(\aleph_1)$ of families $\mathfrak{F}_a$ of sets V to each of which a point of E is interior; there is a point q of E which is contained in a set of each family $\mathfrak{F}_a$.

Theorem 18 still holds if the only sequences considered in property B are those in which there is for each α a point q_a of E which is interior to a set of $\mathfrak{F}_a$ but not contained in any set of $\mathfrak{F}_{a+1}$. This modified form of property B is equivalent (without restrictions on $/E/$) to the property of Theorem 17 regarding ascending sequences.†

From Theorem 12 of "Reducibility," we can secure properties analogous to those of Theorem 18, which are equivalent to separability when the power of E is greater than $\aleph_1$.‡

STATE UNIVERSITY OF IOWA,
IOWA CITY, IOWA.

---

† This equivalence holds if any infinite cardinal μ be substituted for $\aleph_1$. Theorems 17 and 18 hold for any regular cardinal μ. If μ is irregular, the condition of Theorem 17 is only sufficient.

‡ I wish to call attention to related papers by Appert, *Comptes Rendus*, Vol. 194 (1931), p. 2277, and Vol. 196 (1933), pp. 1071-1074; by Haratomi, *Japanese Journal of Mathematics*, Vol. 8 (1931), pp. 113-141; and by myself, "Property C of Hausdorff and the Property of Borel-Lebesgue," soon to appear in the Bulletin of the American Mathematical Society. The functions I and I_0 of section 5, are the complementary set functions of Aumman, *Mathematische Annalen*, Vol. 106 (1932), p. 257.

DECOMPOSITIONS OF CONTINUA BY MEANS OF LOCAL SEPARATING POINTS.

By G. T. Whyburn.

In the first four sections of the present paper there will be developed a method of obtaining decompositions of a continuum M into disjoint subcontinua by means of set-functions satisfying suitable conditions. If these functions are sufficiently restricted, the decompositions obtained will be upper semi-continuous, so that a decomposition space may be defined. The remainder of the paper is devoted to a study of the decompositions obtained when these set-functions are defined in various ways in terms of local separating points of M and cut points of M so that certain ones (and in some cases all) of the conditions considered earlier are satisfied. For example, it will be shown that if M is compact, then for each $p \epsilon M$, there exists a maximal subcontinuum $C_1(p)$ of M containing p and containing only a countable number of local separating points of M; the sets $C_1(p)$ are disjoint and the decomposition of M into sets $C_1(p)$ is upper semi-continuous; the decomposition space C_1 is a regular curve every subcontinuum of which contains uncountably many local separating points both of C_1 and of M. Similarly for each $p \epsilon M$ there exists a maximal subcontinuum $C_3(p)$ of M containing p and such that the local separating points of $C_3(p)$ are countable; furthermore, $C_3(p)$ is identically the sum of all totally imperfect connected subsets of M containing p; this decomposition need not necessarily be upper semi-continuous but it will be in case M is hereditarily locally connected (i. e., if every subcontinuum of M is locally connected), in which case the decomposition space C_3 is hereditarily locally connected and contains no totally imperfect connected subset. Analogous results are obtained using *punctiform* instead of *countable* and likewise by using *cut points* in place of local separating points.

1. *Notation. Conditions on the set-functions.* Unless otherwise specified our space, which we denote by M, will be a compact metric continuum, although it will be seen that a large number of the results are proved without using the compactness of M and hence hold for any metric connected space. We shall use the letter X_0 to designate a countable set, which may vary in the course of a discusssion, and the fact that an expression is set $= X_0$ or is replaced by X_0 means simply that the set represented by this expression is countable, e. g., $A \subset B + X_0$ means that $A - A \cdot B$ is countable or that A is contained in B except possibly for a countable number of points. Similarly

we shall use P_σ to denote a punctiform F_σ which may vary in a discussion. Thus $A = P_\sigma$ means simply that A is a punctiform F_σ, i. e., the sum of a countable number of closed sets no one of which contains a non-degenerate continuum. An arrow will signify implication, e. g., $(d) \rightarrow (c)$ means that condition (d) implies condition (c).

Let $L(C)$ be a set-function defined for all subcontinua C of M. We shall consider two systems of conditions, as follows:

I

(a) $L(C) \subset C + X_0$

(b) $(\bar{E} - E) \cdot L(\bar{E}) = X_0$ (E any connected set)

(c) $C' \cdot L(C) \subset L(C') + X_0$ (C' any subcontinuum of C)

(d) $C' \cdot L(C) \subset L(C')$

(e) $L(C') \subset C' \cdot L(C) + X_0$

(f) $L(C') = C' \cdot L(C) \pm X_0$

II

(α) $L(C) \subset C + P_\sigma$

(β) $(\bar{E} - E) \cdot L(\bar{E}) \subset P_\sigma$ (E any connected set)

(γ) $C' \cdot L(C) \subset L(C\ ) + P_\sigma$ (C' any subcontinuum of C)

(δ) $C' \cdot L(C) \subset L(C')$

(ϵ) $L(C') \subset C' \cdot L(C) + P_\sigma$

(ζ) $L(C') = C' \cdot L(C) \pm P_\sigma$

(η) $L(C)$ is an F_σ.

Between the conditions in system I we have at once the relations:

$$(e) \rightarrow (a),\ (d) \rightarrow (c),\ \text{and}\ [c, e] \rightleftarrows (f).$$

Likewise in system II we have

$$(\epsilon) \rightarrow (\alpha),\ (\delta) \rightarrow (\gamma),\ \text{and}\ [\gamma, \epsilon] \rightleftarrows (\zeta).$$

Finally, it is seen immediately that each condition in system I implies the corresponding one in system II, i. e.,

$$(a) \equiv (\alpha),\ (b) \rightarrow (\beta),\ (c) \rightarrow (\gamma),\ (d) \equiv (\delta),\ (e) \rightarrow (\epsilon)\ \text{and}\ (f) \rightarrow (\zeta).$$

2. *Decompositions with system I.*

(2.1) THEOREM. $[a, b, c]$ *implies that for each* $p \epsilon M$ *there exists a maximal subcontinuum* $C(p)$ *of* M *containing* p *and such that* $L[C(p)]$ *is countable.*

Proof. Let $C(p)$ be the sum of all subcontinua C of M containing p such that $L(C)$ is countable. Let $P = \Sigma p_i$ be a dense subset of $C(p)$. For

each i, let C_i be a continuum $\supset p + p_i$ and such that $L(C_i)$ is countable. Set $C = \sum_1^\infty C_i$. Then $\bar{C} \supset C(p)$. We shall show that $L(\bar{C})$ is countable and hence that $C \equiv C(p)$. We have

$$\begin{aligned} L(\bar{C}) &\subset L(\bar{C})\cdot(\bar{C} - C) + L(\bar{C})\cdot C + X_0 \\ &= L(\bar{C})\cdot(\bar{C} - C) + \Sigma L(\bar{C})\cdot C_i + X_0 \\ &\subset L(\bar{C})\cdot(\bar{C} - C) + \Sigma L(C_i) + X_0 && \text{by } (c) \\ &= X_0 + \Sigma X_0 + X_0 = X_0 && \text{by } (b). \end{aligned}$$

(2. 2) $[a, c]$ *implies that if A and B are continua, $A \cdot B \neq 0$, then $L(A + B) \subset L(A) + L(B) + X_0$.*

For,
$$\begin{aligned} L(A + B) &\subset L(A + B)\cdot A + L(A + B)\cdot B + X_0, && \text{by } (a), \\ &\subset L(A) + X_0 + L(B) + X_0 + X_0, && \text{by } (c), \\ &= L(A) + L(B) + X_0. \end{aligned}$$

(2. 3) COROLLARY. *No two distinct sets $C(p)$ have a common point.*

For if $C(p) \cdot C(q) \neq 0$, then $L[C(p) + C(q)] \subset L[C(p)] + L[C(q)] + X_0 = X_0$.

Whence,
$$C(p) = C(p) + C(q) = C(q).$$

Thus $q \epsilon C(p) \rightarrow C(q) \equiv C(p)$.

(2. 4) LEMMA. *If C is connected, $C \subset M$, there exists a set $X_0 \subset \bar{C} - C$ such that $M - (\bar{C} - C) + X_0$ is connected. Thus no subset of $\bar{C} - C - X_0$ disconnects M.*

Let $\Sigma p_i = P \subset M - \bar{C}$ be dense in $M - \bar{C}$. For each i, let x_i be a point of $\bar{C}$ which is a limit point of the component H_i of $M - \bar{C}$ containing p_i. Set $X_0 = \Sigma x_i$. Then since $C + \Sigma(H_i + x_i)$ is connected and dense in M, no subset of its complement disconnects M. Thus $M - [\bar{C} - C - X_0]$ is connected.

(2. 5) LEMMA. *If $K = Lim\ K_n$, all subcontinua of M, there exists a countable set $X_0 \subset K$ such that $C = K \cdot \Sigma K_n + X_0 + M - K$ is connected. Thus no subset of $K - K \cdot \Sigma K_n - X_0$ disconnects M.**

For each i such that $K_i \cdot K = 0$, let y_i be a limit point in K of the component H_i of $M - K$ containing K_i. Then $\Sigma K_i + \Sigma(H_i + y_i) = D$ is connected and $\bar{D} - D \supset K - K \cdot \Sigma K_n - \Sigma y_i$. Applying (2. 4), we obtain a set

* This lemma is a generalization of a recent result of C. Zarankiewicz, see *Sitzungsberichte der Berliner Mathematischen Gesellschaft*, 1932, p. 43.

X_0 which we may suppose $\supset \Sigma y_i$ and such that $M - (\bar{D} - D - X_0)$ is connected. Since $\bar{D} - D - X_0 \supset K - K \cdot \Sigma K_i - X_0$, we have that $M - K + K \cdot \Sigma K_n + X_0$ is connected.

(2.6) $[b, f]$, $(\rightleftarrows [a, b, c, e])$, *implies that the decomposition of M into sets $C(p)$ is upper semi-continuous.**

Proof. Let $K_n = C(p_n)$, $(n = 1, 2, 3, \cdots)$, be any sequence of sets $C(p)$ converging to a limit continuum K. Applying (2.5) we obtain a connected set C such that $\bar{C} = M$ and

$$\text{(i)} \qquad K \subset \Sigma K_n + X_0 + (\bar{C} - C).$$

Let $p \epsilon K$. Then by (2.2) we have

$$\text{(ii)} \qquad L[C(p) + K] \subset L[C(p)] + L(K) + X_0 = L(K) + X_0,$$

since $L[C(p)] = X_0$. Now

$$\text{(iii)} \qquad L(K) \subset L(M) \cdot K + X_0, \qquad \text{by } (e),$$

and

$$\begin{aligned} \text{(iv)} \qquad L(M) \cdot K &\subset \Sigma L(M) \cdot K_n + X_0 \cdot L(M) + L(M) \cdot (\bar{C} - C) && \text{by (i)},\\ &\subset \Sigma L(K_n) + \Sigma X_0 + X_0 + X_0 && \text{by } (c) \text{ and } (b),\\ &= X_0. \end{aligned}$$

Thus, by (iii), $L(K) = X_0$. Whence, by (ii), $L[C(p) + K] = X_0$, which proves $K \subset C(p)$.

(2.7) $[a, b]$ *implies that if K is any continuum of convergence of M, $L(M) \cdot K = X_0$.*

Let $K = \text{Lim } K_n$, where $K_n \cdot K = 0$ for each n. By (2.5) we have a set X_0 such that $C = M - K + X_0$ is connected. Now $L(M) \cdot K \subset L(M) \cdot (\bar{C} - C) \cdot K + L(M) \cdot K \cdot C + X_0 \subset X_0 + X_0 + X_0 = X_0$, since $M = \bar{C}$ and $K \cdot C = X_0$.

(2.8) $[b, f]$ *implies that if K is any continuum of convergence of M and $p \epsilon K$, then $K \subset C(p)$.*

By virtue of the definition of $C(p)$ we have only to show that $L[C(p) + K] = X_0$. Now by (2.2) we have

* That is, the collection of sets $[C(p)]$ is upper semi-continuous. This means that if $C(p_1)$, $C(p_2)$, $\cdots$ is any convergent sequence of these sets, then there exists some single set $C(p)$ which contains $\text{Lim } [C(p_i)]$. See R. L. Moore, "Foundations of point set theory," *American Mathematical Society colloquium publications* (1932), Ch. V; also P. Alexandroff, *Mathematische Annalen*, Vol. 96, pp. 555-571.

$$\begin{aligned} L[C(p)+K] &\subset L(K)+L[C(p)]+X_0 \\ &\subset L(M)\cdot K+X_0, \text{ by } (f) \text{ and since } L[C(p)]=X_0 \\ &\subset X_0 \qquad \text{by } (2.7). \end{aligned}$$

(2.9) *If C denotes the hyperspace * of the decomposition into sets $C(p)$ given by a function L satisfying $[b, f]$, then C is hereditarily locally connected.*

For suppose C has a continuum of convergence K. Then $\hat{K}$ † is a continuum of convergence of M and by (2.8) we have $\hat{K} \subset C(p)$ for any $p \epsilon \hat{K}$. But $\hat{K}$ is the sum of a certain collection of sets $C(p)$; whence $\hat{K} = C(p)$. Thus K is a single point in C, contrary to supposition.

3. *Decompositions with system II.*

(3.1) THEOREM. $[\alpha, \beta, \delta]$ *or* $[\alpha, \beta, \gamma, \eta]$ *implies that for each* $p \epsilon M$ *there exists a maximal subcontinuum* $D(p)$ *of* M *containing* p *and such that* $L[D(p)]$ *is punctiform.*

Proof. Let $D(p)$ be the sum of all subcontinua D of M containing p and such that $L(D)$ is punctiform. Let $P = \Sigma p_i$ be dense in $D(p)$ and, for each i, let D_i be a subcontinuum of M containing $p + p_i$ and such that $L(D_i)$ is punctiform. Clearly if $D = \Sigma D_i$, we have $\bar{D} \supset D(p)$. We have, to show that $L(\bar{D})$ is punctiform. If, on the contrary, it contained a continuum K we would have

(i) $$K = \Sigma K \cdot D_i + K \cdot (\bar{D} - D) + P_\sigma.$$

By (β),

(ii) $$K \cdot (\bar{D} - D) \subset L(D) \cdot (\bar{D} - D) \subset P_\sigma.$$

On the other hand (δ) gives

(iii) $$K \cdot D_i \subset L(\bar{D}) \cdot D_i \subset L(D_i) \subset P_\sigma$$

while $[\gamma, \eta]$ gives

(iv) $$K \cdot D_i \subset L(\bar{D}) \cdot D_i \subset L(D_i) + P_\sigma = P_\sigma + P_\sigma = P_\sigma.$$

Now either (iii) or (iv) gives

(v) $$\Sigma K \cdot D_i \subset \Sigma P_\sigma = P_\sigma$$

---

* That is, the space whose elements are the sets $C(p)$ and in which distance has been defined as, e.g., $\rho[C(p), C(q)] = \min [\rho(x, y)]$, $x \epsilon C(p)$, $y \epsilon C(q)$. It is known that since the decomposition is upper semi-continuous, this space will be metric, compact and connected. See Moore, *loc. cit.*, and Alexandroff, *loc. cit.* We shall call a space obtained in this way by a decomposition of M the *decomposition space* or the *hyperspace of the decomposition.*

† If X is a set of elements in a decomposition space C, $\hat{X}$ will denote the point set in M obtained by adding together all the elements of X.

and (i), (ii), (v) give $K \subset P_\sigma$, which is impossible.

(3. 2) $[\alpha, \gamma]$ *implies that if* A *and* B *are continua,* $A \cdot B \neq 0$ *then* $L(A+B) \subset L(A) + L(B) + P_\sigma$.

For
$$\begin{aligned} L(A+B) &\subset A \cdot L(A+B) + B \cdot L(A+B) + P_\sigma \\ &\subset L(A) + P_\sigma + L(B) + P_\sigma + P_\sigma, \text{ by } (\gamma), \\ &= L(A) + L(B) + P_\sigma. \end{aligned}$$

(3. 3) COROLLARY. *No two distinct sets* $D(p)$ *have a common point.*

(3. 4) THEOREM. $[\alpha, \beta, \delta, \epsilon]$ *or* $[\alpha, \beta, \gamma, \eta, \epsilon]$ *implies that the decomposition of* M *into sets* $D(p)$ *is upper semi-continuous.*

Proof. Let $K_n = D(p_n)$, $n = 1, 2, 3, \cdots$, be any sequence of sets $D(p)$ converging to a limit continuum K. Applying (2. 5) we obtain a connected set C such that $\bar{C} = M$ and

(i) $$K \subset \Sigma K_n + X_0 + (\bar{C} - C).$$

Let $p \epsilon K$. Then by (3. 2) we have

(ii) $$L[D(p) + K] \subset L(K) + L[D(p)] + P_\sigma,$$

Now let us suppose $L[C(p) + K]$ contains a continuum N. Then (ii) and (η) would give that $L(K)$ contains a continuum, while (ii) and (δ) would give that $N \cdot D(p) \subset L[D(p)]$, which is punctiform, so that $K \cdot N \subset L(K)$ cannot be punctiform. Thus in either case $L(K)$ contains a continuum N_1. Now

(iii) $$L(K) \subset L(M) \cdot K + P_\sigma, \qquad \text{by } (\epsilon),$$

and

(iv) $$\begin{aligned} L(M) \cdot K &\subset \Sigma L(M) \cdot K_n + X_0 \cdot L(M) + L(M) \cdot (\bar{C} - C), \quad \text{by (i)}, \\ &\subset \Sigma L(K_n) + \Sigma P_\sigma + X_0 + P_\sigma \qquad (\text{by } \gamma) \\ &= \Sigma L(K_n) + P_\sigma. \end{aligned}$$

Now (η) together with (iv) would give

$$L(K) \subset \Sigma P_\sigma + P_\sigma = P_\sigma,$$

which is impossible since $L(K) \supset N_1$. On the other hand (δ) with (iv) gives

$$N_1 \subset \Sigma N_1 L(K_n) + P_\sigma;\ P_\sigma = N_1 \cdot K_n \subset N_1 \cdot L(K_n) = \Sigma N_1 \cdot K_n + P_\sigma$$

Thus

$$N_1 \subset P_\sigma + P_\sigma = P_\sigma,$$

which is impossible. Thus the supposition that $L[D(p) + K]$ is not punctiform leads to a contradiction. Therefore $K \subset D(p)$, which proves our theorem.

(3. 5) $[\alpha, \beta]$ *implies that if* K *is a continuum of convergence of* M, $L(M) \cdot K \subset P_\sigma$.

Let $K = \operatorname{Lim} K_n$, where $K \cdot K_n = 0$, $n = 1, 2, 3, \cdots$. By (2. 5) we have a set X_0 such that $C = M - K + X_0$ is connected. Now

$$\begin{aligned} L(M) \cdot K &\subset L(M) \cdot (\bar{C} - C) \cdot K + L(M) \cdot K \cdot C + P_\sigma, \qquad \text{by } [\alpha, \beta] \\ &\subset P_\sigma + X_0 + P_\sigma = P_\sigma, \end{aligned}$$

since $M = \bar{C}$ and $K \cdot C = X_0$.

(3. 6) $[\beta, \delta, \epsilon]$ *or* $[\beta, \gamma, \eta, \epsilon]$ *implies that if* K *is any continuum of convergence and* $p \epsilon K$, *then* $K \subset D(p)$.

Clearly we have only to show that $L[K + D(p)]$ is punctiform. Now by (3. 2) we have

$$\begin{aligned} L[K + D(p)] &\subset L(K) + L[D(p)] + P_\sigma \\ &\subset L(M) \cdot K + L[D(p)] + P_\sigma \qquad \text{by } (\epsilon) \\ &\subset L[D(p)] + P_\sigma \qquad \text{by } (3.\ 5). \end{aligned}$$

Thus since $L[D(p)]$ is punctiform, (η) would give at once that $L[K + D(p)] \subset P_\sigma$ and hence is punctiform. On the other hand, if we suppose that $L[K + D(p)]$ contains a continuum N, we have from the above

$$N \subset N \cdot L[D(p)] + P_\sigma$$

and (δ) would give $N \cdot L[D(p)] \equiv N \cdot D(p) = P_\sigma$, since $L[D(p)]$ is punctiform. Thus $N \subset P_\sigma$, which is impossible. Thus in either case $L[K + D(p)]$ is punctiform.

(3. 7) *If* D *denotes the hyperspace of the decomposition of* M *into sets* $D(p)$ *given by a function* L *satisfying either* $[\beta, \delta, \epsilon]$ *or* $[\beta, \gamma, \eta, \epsilon]$, *then* D *is hereditarily locally connected.*

The proof is identical with the proof of (2. 9), substituting D for C.

4. *The case where* M *is hereditarily locally connected.* It is to be noted that while the existence of the sets $C(p)$ and $D(p)$ was established above on the basis of $[a, b, c]$ and $[\alpha, \beta, \delta]$ or $[\alpha, \beta, \gamma, \eta]$, respectively, an extra condition (e) or (ϵ) was added in order to establish the upper semi-continuity of the decompositions of M into these respective sets. It will be seen from the examples given below in § 7 that in the general case this extra condition is not redundant. However, we proceed now to show that in case M is hereditarily locally connected, the upper semi-continuity of the decomposition into sets $C(p)$ and $D(p)$ results from $[a, b, c]$ and $[\alpha, \beta, \delta]$ or $[\alpha, \beta, \gamma, \eta]$, respectively, without supposing the extra conditions (e) or (ϵ).

(4.1) LEMMA. *If H is a hereditarily locally connected compact continuum and G is any collection of disjoint continua filling up H, then in order that G be upper semi-continuous it is necessary and sufficient that the sum of no countable number (>1) of elements of G be a connected point set.**

Proof. The condition is necessary. For if G is upper semi-continuous, distance can be so defined between the elements of G that the resulting space is metric. Since no countable set of points in a metric space can be connected, it follows that no countable set of elements of G can be a connected set of elements and hence the point set obtained by adding together any countable number (>1) of elements of G cannot be connected.

The condition is also sufficient. For if G is not upper semi-continuous, there must exist a number $d>0$ and an infinite set X of elements of G each of diameter $>d$. Then X contains a convergent sequence $X_1, X_2, \cdots$ of elements converging to a limit continuum K. Now since H is hereditarily locally connected, K cannot be a continuum of convergence of H. Thus for infinitely many i's, say $i_1, i_2, \cdots$, we have $X_{i_n}\cdot K \neq 0$ for each $n=1,2,\cdots$ But then it is seen at once that $\sum_{n=1}^{\infty} X_{i_n}$ is connected.

(4.2) *If M is hereditarily locally connected, the decomposition of M into sets $C(p)$ by any function L satisfying $[a, b, c]$ is upper semi-continuous.*

For if not, then by (4.1) there exists a countable sequence $C(p_1)$, $C(p_2)$, $C(p_3), \cdots$ of distinct sets $C(p)$ such that $C = \Sigma C(p_i)$ is connected. But

$$\begin{aligned} L(\bar{C}) &\subset L(\bar{C})\cdot(\bar{C}-C) + L(\bar{C})\cdot C + X_0 && \text{by } (a)\\ &\subset X_0 + \Sigma L(\bar{C})\cdot C(p_i) + X_0 && \text{by } (b)\\ &\subset \Sigma[L[C(p_i)] + X_0] + X_0 && \text{by } (c)\\ &= \Sigma X_0 + X_0 = X_0, \text{ since } L[C(p_i)] = X_0. \end{aligned}$$

This is impossible, since $\bar{C} \supset C(p_1)$, and $C(p_1)$ is the maximal subcontinuum C of M containing p such that $L(C) = X_0$.

---

* That, contrary to a statement of Kuratowski (*Fund. Math.*, Vol. 11 (1928), p. 180) not every decomposition of such a continuum H is necessarily upper semi-continuous is seen from the following example. Let H denote that part of the continuum ϕ described by the author in *Mathematische Annalen*, Vol. 102, p. 333, for which $\sqrt{2}/20 \leq z \leq 1-\sqrt{2}/20$. Then H is a hereditarily locally connected continuum. Let G denote the collection of continua whose elements are the continua $H\cdot C_n$ together with all points of $H - \Sigma H\cdot C_n$. Then G fills up H and the elements of G are disjoint continua; but G is not upper semi-continuous because Lim $[H\cdot C_n]$ contains both of the points $(0, 0, \sqrt{2}/20)$ and $(0, 0, 1-\sqrt{2}/20)$, and each of these points is an element of G.

COROLLARY. *In the general case, the sum of no countable number* (> 1) *of sets* $C(p)$ *can be connected.*

(4.3) *If M is hereditarily locally connected, the decomposition of M into sets $D(p)$ by any function L satisfying either* $[\alpha, \beta, \delta]$ *or* $[\alpha, \beta, \gamma, \eta]$ *is upper semi-continuous.*

For if not, then just as above there exists a countable sequence $D(p_1)$, $D(p_2), \cdots$ of distinct sets $D(p)$ whose sum D is connected. Then

$$\begin{aligned} L(\bar{D}) &\subset L(\bar{D}) \cdot (\bar{D} - D) + L(\bar{D}) \cdot D + P_\sigma && \text{by } (\alpha) \\ &\subset L(\bar{D}) \cdot \Sigma D(p_i) + P_\sigma && \text{by } (\beta) \\ &\subset \Sigma L[D(p_i)] + P_\sigma && \text{by } (\gamma) \end{aligned}$$

Now (η) would give at once $L(\bar{D}) \subset \Sigma P_\sigma + P_\sigma = P_\sigma$, since each $L[D(p_i)]$ is punctiform. On the other hand, if $L(\bar{D})$ contains a continuum N, (δ) would give $N \cdot D(p_i) = N \cdot L[D(p_i)] \cdot D(p_i)$, for each i. Thus since each $L[D(p_i)]$ is punctiform, $P_\sigma = N \cdot D = \Sigma N \cdot L[D(p_i)] \cdot D(p_i)$. So that $N \subset N \cdot D + N(\bar{D} - D) + P_\sigma \subset P_\sigma$ which is impossible. Thus in either case it follows that $L(\bar{D})$ is punctiform, contrary to the fact that $\bar{D}$ contains more than one set $D(p)$.

COROLLARY. *If M is hereditarily locally connected, for any $\epsilon > 0$, at most a finite number of the sets $C(p)$ or $D(p)$ are of diameter $> \epsilon$.*

5. *Applications: some particular functions L.* We define four functions L as follows:

$L_1(C)$ = the set of all local separating points of M belonging to C.
$L_2(C)$ = the set of all cut points of M belonging to C.
$L_3(C)$ = the set of all local separating points of C.
$L_4(C)$ = the set of all cut points of C.

(5.1) THEOREM. *The functions $L_1(C)$ and $L_2(C)$ satisfy* $[a, b, c, d, e, f]$ *and* $[\alpha, \beta, \gamma, \delta, \epsilon, \zeta]$.

It suffices to prove $[b, d, f]$ since, by § 1, this combination implies all the other conditions. To prove (b) it suffices, since $L_2(\bar{E}) \subset L_1(\bar{E})$, to show that if E is any connected subset of M, $L_1(\bar{E}) \cdot (\bar{E} - E) = X_0$. If we suppose, on the contrary, that $L_1(\bar{E}) \cdot (\bar{E} - E)$ is uncountable, then since every point of this set is a local separating point of M it follows by a theorem of the author's * that there exists a point x of this set which is separated in $\bar{E}$

* See *Monatshefte für Mathematik und Physik*, Vol. 36 (1929), pp. 309-311.

from some point of E by two points z and w of this set. But clearly this is impossible, since $E + x$ is connected. Now by definition we have

$$L_i(C) = L_i(M) \cdot C,\ L_i(C') = L_i(M) \cdot C', \qquad (i = 1, 2).$$

Thus $C' \subset C$ gives $C' \cdot L_i(C) \equiv L_i(M) \cdot C' \equiv L(C')$, which implies $[d, f]$.

(5.2) *The functions $L_3(C)$ and $L_4(C)$ satisfy $[a, b, c]$ and $[\alpha, \beta, \gamma]$.*

By definition we have $L_4(C) \subset L_3(C) \subset C$ which gives (a). To prove (b), let E be any connected subset of M. Since $L_4(\bar{E}) \subset L_3(\bar{E})$, we have only to show that $L_3(\bar{E}) \cdot (\bar{E} - E) = H$ is countable. If not, then since every point of H is a local separating point of $\bar{E}$ it follows just as in the proof of (5.1) that there exists a point x of H which can be separated in $\bar{E}$ from a point of E by the removal of two points of H. But clearly this is impossible, since $E + x$ is connected and contains no point of H.

That $L_4(C)$ satisfies (c) is equivalent to the well known theorem * that if $C' \subset C$, then all save a countable number of the cut points of C that are on C' are cut points also of C'. To show that $L_3(C)$ also satisfies (c) let us suppose on the contrary that C' contains an uncountable set H of local separating points of C which are not local separating points of C'. But by the theorem of the author's quoted above, H contains a point x which is a point of order 2 in C relative to H; and it is seen at once that since x locally separates C it must also locally separate C', contrary to supposition.

Thus we have shown that $L_3(C)$ and $L_4(C)$ satisfy $[a, b, c]$ and since $[a, b, c] \rightarrow [\alpha, \beta, \gamma]$, (5.2) is established.

(5.3) *If M is locally connected, $L_1(C)$ and $L_2(C)$ satisfy also (η); and if M is hereditarily locally connected, $L_3(C)$ and $L_4(C)$ satisfy (η).*

It is to be seen at once that this is equivalent to the well known facts that the totality of all cut points † and of all local separating points ‡ respectively of a locally connected continuum are F_σ sets.

As a consequence of (5.1) and (5.2) we have by (2.1) and (3.1) that the functions L_1, L_2, L_3, L_4 yield decompositions of M into sets $C_1(p)$, $C_2(p)$, $C_3(p)$, $C_4(p)$ respectively, L_1 and L_2 yield decompositions of M into sets $D_1(p)$, $D_2(p)$ respectively, and in case M is hereditarily locally connected L_3 and L_4 yield decompositions of M into sets $D_3(p)$ and $D_4(p)$, where $C_i(p)$ $(i = 1, 2, 3, 4)$ is the maximal subcontinuum C of M containing p

* See R. L. Moore, *Proceedings of the National Academy of Sciences*, Vol. 9 (1923), pp. 101-106.

† See Zarankiewicz, *Fundamenta Mathematicae*, Vol. 9 (1927), pp. 124-171.

‡ See Whyburn, *Mathematische Annalen*, Vol. 102 (1930), p. 318.

and such that $L_i(C)$ is countable and $D_i(p)$ $(i=1,2,3,4)$ is the maximal subcontinuum D of M containing p and such that $L_i(D)$ is punctiform. Thus we have

(5.4) THEOREM. *For each $p \epsilon M$ there exists maximal subcontinua $C_1(p)$, $C_2(p)$, $C_3(p)$, $C_4(p)$, $D_1(p)$, $D_2(p)$ and, in case M is hereditarily locally connected, $D_3(p)$ and $D_4(p)$, containing p and such that*

$C_1(p)$ *contains at most a countable number of local separating points of M*
$C_2(p)$ *contains at most a countable number of cut points of M*
$C_3(p)$ *has at most a countable number of local separating points*
$C_4(p)$ *has at most a countable number of cut points*
$D_1(p)$ *contains only a punctiform set of the local separating points of M*
$D_2(p)$ *contains only a punctiform set of the cut points of M*
$D_3(p)$ *has only a punctiform set of local separating points*
$D_4(p)$ *has only a punctiform set of cut points.*

By virtue of (2.6), (3.4), and (5.1) we have that the decompositions of M into sets $C_1(p)$, $C_2(p)$, $D_1(p)$, $D_2(p)$ are upper semi-continuous; and if M is hereditarily locally connected, (4.2), (4.3), (5.2), (5.3) yield that the decompositions of M into sets $C_3(p)$, $C_4(p)$, $D_3(p)$, $D_4(p)$ likewise are upper semi-continuous. Let us denote the decomposition spaces by C_1, C_2, C_3, C_4, D_1, D_2, D_3, D_4 respectively. We thus have

(5.5) THEOREM. *The decompositions of M into sets $C_i(p)$ and $D_i(p)$, $(i=1,2)$, are upper semi-continuous; and if M is hereditarily locally connected, so are the decompositions into sets $C_i(p)$ and $D_i(p)$, $(i=3,4)$. Furthermore all of the decomposition spaces C_i and D_i $(i=1,2,3,4)$ are hereditarily locally connected continua.*

A detailed study of these decomposition spaces will be made below in § 6.

Regarding the inclusion relations among the sets $C_i(p)$ and $D_i(p)$ we have

(5.6) *For each $p \epsilon M$,* (i) $C_3(p) \subset C_1(p) \subset C_2(p)$, (ii) $C_3(p) \subset C_4(p) \subset C_2(p)$. *In general $C_1(p)$ and $C_4(p)$ are independent, but in case M is locally connected, $C_2(p) \equiv C_4(p)$ which, together with* (i), *gives $C_1(p) \subset C_4(p)$.*

Relations (i) and (ii) result immediately from the fact that any cut point of a continuum is also a local separating point. The equality of $C_2(p)$ and $C_4(p)$ in case M is locally connected follows from the fact that in this

case $C_2(p)$ is an A-set * in M and that any cut point of an A-set is a cut point also of M. Similarly we have

(5.7) *For each* $p\epsilon M$, (i) $D_1(p) \subset D_2(p), C_1(p) \subset D_1(p), C_2(p) \subset D_2(p)$; *and in case M is hereditarily locally connected,* (ii) $D_3(p) \subset D_1(p) \subset D_2(p) \equiv D_4(p)$, *and* $C_3(p) \subset D_3(p)$, $C_4(p) \subset D_4(p)$.

The relations (i) and the inclusion $D_1(p) \subset D_2(p)$ of (ii) follow from the fact that every cut point is also a local separating point. To prove the inclusion $D_3(p) \subset D_1(p)$ of (ii) we suppose, on the contrary, that $D_3(p)$ contains a continuum K of local separating points of M. Then K contains an arc ab and † either every inner point of ab separates a and b in $D_3(p)$ or ab contains an arc st which is free in some cyclic element of $D_3(p)$; and in either case ab contains an arc of local separating points of $D_3(p)$, contrary to the fact that $L_3[D_3(p)]$ is punctiform.

The identity of $D_2(p)$ and $D_4(p)$ results just as in the proof of (5.6) from the fact that $D_2(p)$ is an A-set. To see this, let E be any non-degenerate cyclic element of M which contains at least one point of $D_2(p)$. Then since M is locally connected, we have $L_2[D_2(p)] = P_\sigma$ and $L_2(E) = X_0$, so that $L_2[D_2(p) + E] = P_\sigma + X_0 = P_\sigma$, which gives $E \subset D_2(p)$. Thus $D_2(p)$ is an A-set, and since every cut point of an A-set is a cut point also of M, we have $D_2(p) \equiv D_4(p)$. The remaining relations (ii) are trivial.

Since all save a countable number of the local separating points (hence also of the cut points) of M are points of order 2 of M, it follows that in any decomposition of M into disjoint continua at most a countable number of the non-degenerate elements can contain local separating points or cut points of M. This fact together with the fact that the sets $C_1(p)$ and $C_3(p)$ contain at most a countable number of cut points of M gives

(5.8) *At most a countable number of the sets $C_i(p)$ and $D_i(p)$ $(i = 1, 2, 3, 4)$ can contain local separating (or cut) points of M. For all save a countable number of local separating points p of M we have $C_1(p) = C_3(p) = p$, and for all save a countable number of cut points of M we have $C_2(p) = C_4(p) = p$.*

We conclude this section by giving an alternate method of obtaining the sets $C_1(p)$ and $D_1(p)$. Let X denote the set of all points $x\epsilon M$ such that every subcontinuum of M containing x contains uncountably many local separating points of M. Similarly let Y denote the set of all points $y\epsilon M$

* That is, a subcontinuum of M which is a sum of cyclic elements of M. See Kuratowski and Whyburn, *Fundamenta Mathematicae*, Vol. 16 (1930), p. 309.

† See my paper, *American Journal of Mathematics*, Vol. 55 (1933), p. 148.

such that every subcontinuum of M containing y contains a continuum of local separating points. Then we have

(5.9) *The sets $C_1(p)$ are exactly the points of X together with the components of $M - X$, i. e., for $p \epsilon X$, $C_1(p) = p$, while for $p \epsilon (M - X)$, $C_1(p)$ is the component of $M - X$ containing p. Likewise the sets $D_1(p)$ are the points of Y and the components of $M - Y$.*

If $p \epsilon X$, then since $C_1(p)$ contains only a countable number of local separating points of M we have $C_1(p) = p$. If $p \epsilon (M - X)$ and K denotes the component of $M - X$ containing p, then since X contains all save a countable number of the local separating points $L_1(M)$, and $(\bar{K} - K) \cdot L_1(M) = X_0$, we have $L_1(M) \cdot \bar{K} = X_0$ and hence $\bar{K} = K$ and $K \subset C_1(p)$. Since any larger continuum containing K contains a point of X, it would contain uncountably many points of $L_1(M)$ and hence could not be $C_1(p)$. Thus $K \equiv C_1(p)$.

A similar argument establishes the latter part of the theorem.

It is obvious that we could give similar alternate definitions for the sets $C_2(p)$ and $D_2(p)$, using the cut points of M instead of the local separating points.

6. *The decomposition spaces C_i, D_i $(i = 1, 2, 3, 4)$.*

(6.1) *Under any upper semi-continuous decomposition of M into disjoint continua with hyperspace H, if for a given $p \epsilon M$ we have also $p \epsilon H$, then* (i) *p is a cut point of H if and only if it is a cut point of M,* (ii) *p is a local separating point of H if and only if it is a local separating point of M.*

The truth of (i) results immediately from the fact * that under the given conditions a set K of elements of H is connected if and only if the set K of points in M is connected.

Likewise (ii) follows from similar considerations. For if p is a local separating point of M, there exists a neighborhood R of p in M such that p separates $\bar{R}$ between some pair of points on the component of $\bar{R}$ containing p. Since $p \epsilon H$ there exists a neighborhood G of p in H with $\hat{\bar{G}} \subset R$. Since p is an interior point of $\hat{G}$ (rel. M), we have at once $\hat{\bar{G}} - p = \hat{G}_1 + \hat{G}_2$, where $\hat{G}_1$ and $\hat{G}_2$ are mutually separated and each intersects the component $\hat{C}$ of $\hat{\bar{G}}$ containing p. Then $\bar{G} - p = G_1 + G_2$, G_1 and G_2 are mutually separated and each intersects C, the component of $\bar{G}$ containing p, and thus p locally separates H. On the other hand, if p is a local separating point of H, we

* See R. L. Moore, *Foundations of Point Set Theory*, Ch. V.

have for some neighborhood G of p in H, $\bar{G} - p = G_1 + G_2$, where G_1 and G_2 are mutually separated and each intersects the component C of $\bar{G}$ containing p. And since $\hat{G}$ is open in M and $\hat{G}_1$ and $\hat{G}_2$ are mutually separated, it is clear that p is a local separating point of M.

(6.2) THEOREM. *For any compact continuum M: (1) C_1 is a regular curve every subcontinuum of which contains uncountably many local separating points of C_1 (also of M); (2) C_2 is a dendrite* * *(every subcontinuum contains uncountably many cut points of M); (3) D_1 is a regular curve every subcontinuum of which contains a continuum of local separating points of D_1, and thus no cyclic element of D_1 has a continuum of condensation; (4) D_2 is a dendrite (every subcontinuum contains a continuum of cut points of M). If M is hereditarily locally connected, then: (5) C_3 is hereditarily locally connected continuum every subcontinuum of which has uncountably many local separating points; (6) $C_4 \equiv C_2$; (7) D_3 is a hereditarily locally connected continuum every subcontinuum of which has a continuum of local separating points; (8) $D_4 \equiv D_2$. Furthermore each of these properties is characteristic for the respective decompositions in the sense that if M has one of these properties to begin with, the corresponding decomposition is trivial, giving merely the points of M.*

Proof. To prove (1), let K be any subcontinuum of C_1. Then since $\hat{K}$ is a subcontinuum of M containing more than one set $C_1(p)$, $\hat{K}$ contains an uncountable set H of local separating points of M. By (5.8), H contains an uncountable set U such that for each $p \epsilon U$, $C_1(p) = p$. Whence $p \epsilon C_1$, $U \subset K$ and, by (6.1), each $p \epsilon U$ is a local separating point of C_1. That C_1 is a regular curve results from the property just proved.

(2) is proved in exactly the same manner. For if K is any subcontinuum of C_2, $\hat{K}$ contains an uncountable set H of cut points of M, and H contains an uncountable set U such that for each $p \epsilon U$, $C_2(p) = p$. Whence $p \epsilon C_2$, $U \subset K$ and, by (6.1), each $p \epsilon U$ is a cut point of C_2. Therefore † C_2 is a dendrite.

To prove (3), let K be any subcontinuum of D_1. Then $\hat{K}$ contains a continuum N of local separating points of M. Let H be the continuum in

* I. e., a locally connected continuum containing no simple closed curve.

† See R. L. Moore, *Proceedings of the National Academy of Sciences, loc. cit.* It is to be noted that the decomposition space C_2 is identical with that obtained by R. L. Moore by another method of approach (See Moore, *Foundations of Point Set Theory*, p. 342). In other words, the sets $C_2(p)$ could be defined or characterized, as is done by Moore, as the point p together with all points x of M which are not separated in M from p by each point of an uncountable set of points of M.

D_1 consisting of all elements in D_1 which intersect N. It follows by (5.8) that the non-degenerate elements in H are countable and hence, by (6.1), that the non-local separating points of D_1 on H are countable. Since the set of non-local separating points of D_1 on H is a G_δ, it cannot be dense on H. Hence H, and therefore K, contains a continuum * every point of which is a local separating point of D_1. It follows at once that every subcontinuum of a cyclic element X of D_1 contains an arc that is free in X, so that X has no continuum of condensation.

The proof of (4) so closely parallels that of (3) that we do not give it.

To prove (5), let E be any subcontinuum of C_3. Since $\hat{E}$ is a subcontinuum of M containing more than one set $C_3(p)$, $\hat{E}$ has an uncountable set U of local separating points. Now at most a countable number of the non-degenerate elements of C_3 in E can contain points of U, and each of these contains only a countable number of points of U. Thus all save a countable number of points of U are elements of C_3 and, by (6.1), each of these is a local separating point of E.

Parts (6) and (8) follow from (5.6) and (5.7) respectively.

Finally, (7) follows by a combination of the methods of arguments used in the proofs of parts (3) and (5) which is sufficiently obvious to be omitted.

The fact that the spaces C_i and D_i possess the characteristic properties just established makes a further study of continua having these respective properties highly desirable. Some of these will be considered later in § 8. For the present we consider continua M having the property, just proved for C_1, that every subcontinuum contains uncountably many local separating points of M. We have

(6.3) *If a continuum M has the property that each of its subcontinua contains uncountably many local separating points of M, then:* (i) *every connected subset of M contains uncountably many local separating points of M;* (ii) *the dimension (Menger-Urysohn) of the set of ramification points of M is 0;* (iii) *for each connected subset G of M, dim $(\bar{G} - G) = 0$;* (iv) *M is disconnected by the omission of any non-punctiform subset S such that $\overline{M - S} \supset S$;* (v) *$M$ is a regular curve such that every irreducible cutting of M between any two of its points is punctiform.*

If use is made of the result of the author's † that any subset of a regular curve of dimension > 0 contains a connected set, the proof of (6.3) will present no difficulties and hence it is omitted.

---

* See my paper in *Mathematische Annalen, loc. cit.*, pp. 318-319.

† See *American Journal of Mathematics*, Vol. 53 (1931), p. 379.

Note: In another paper * we have shown that in order for every connected subset of a continuum M to be a G_δ it is necessary and sufficient that the non-local-separating points of M be countable. This condition is readily seen to be equivalent to the condition that for each connected subset G of M, the set $\bar{G} - G$ be countable. Also from (2.4), § 2, it follows that this is equivalent to the condition that M be disconnected by the omission of every uncountable subset U such that $\overline{M - U} \supset U$. Thus, analogous to (6.3), we have

(6.4) *For any continuum M the following properties are equivalent:* (i) *that every connected subset be a G_δ;* (ii) *that the non-local-separating points of M be countable;* (iii) *that for every connected subset G of M, $\bar{G} - G$ be countable;* (iv) *that M be disconnected by the omission of any uncountable subset U such that $\overline{M - U} \supset U$;* (v) *that M be a regular curve such that every irreducible cutting of M between any two points is countable.*

7. *Examples.*

(7.1) Let T denote the Sierpinski triangle curve.† For each $p \epsilon T$ we have $C_1(p) = C_2(p) = C_3(p) = C_4(p) = D_1(p) = D_2(p) = D_3(p) = D_4(p) = T$. Consequently every one of the decomposition spaces C_i and D_i $(i = 1, 2, 3, 4)$ reduces to a single point.

(7.2) Let $E = I + S$, where S is that part of the graph of $y = \sin 1/x$ for which $0 < x \leq 1$ and I is the limiting continuum (i. e., the interval $(-1, 1)$ of the y-axis) of S. Then for $p \epsilon S$, $C_1(p) = C_2(p) = C_3(p) = C_4(p) = D_1(p) = D_2(p) = D_3(p) = D_4(p) = p$. While for $p \epsilon I$, $C_1(p) = C_2(p) = D_1(p) = D_2(p) = I$; $C_3(p) = C_4(p) = D_3(p) = D_4(p) = p$. Consequently the spaces C_1, C_2, D_1, and D_2 are simple arcs and the remaining decompositions are not upper semi-continuous.

(7.3) Let $F = S + \Delta$, where S is the same as above and Δ is a triangle having I as one side and lying otherwise to the left of the y-axis. In this case it is readily seen that the spaces C_1 and D_1 are each homeomorphic to the point set obtained by adding a circle and a straight line interval which has one of its end points and only this on the circle.

(7.4) Let $H = T + I + S$, where I has just one point in common with T and $T \cdot S = 0$. Then for $p \epsilon S$, $C_1(p) = C_2(p) = C_3(p) = C_4(p) = D_1(p) = D_2(p) = D_3(p) = D_4(p) = p$. For $p \epsilon (I - x)$, $C_1(p) = C_2(p) = D_1(p)$

* *Bulletin of the American Mathematical Society*, Vol. 38 (1933), p. 98.

† See Sierpinski, *Comptus Rendus*, Vol. 162, p. 629.

$= D_2(p) = T + I$; $C_3(p) = C_4(p) = D_3(p) = D_4(p) = p$. For $p \epsilon T$, $C_1(p) = C_2(p) = D_1(p) = D_2(p) = T + I$; $C_3(p) = C_4(p) = D_3(p) = D_4(p) = T$. Therefore the spaces C_1, C_2, D_1, D_2 are all simple arcs, while the remaining decompositions are not upper semi-continuous.

(7.5) Let W denote the curve obtained by taking an equilateral triangle Δ with base B and joining the mid points of the two sides to the mid point of B by intervals, then taking the two smaller equilateral triangles thus formed and joining the mid points of their sides to the mid points of their respective bases (on B), and so on indefinitely. Then for $p \epsilon W$, $C_2(p) = D_2(p) = C_4(p) = D_4(p) = W$; $C_3(p) = D_3(p) = p$. For $p \epsilon B$, $C_1(p) = D_1(p) = B$. For $p \epsilon (W - B)$, $C_1(p) = D_1(p) = p$. Thus the spaces C_2, D_2, C_4, D_4 reduce to single points; C_3 and D_3 are identical with W, while C_1 and D_1 are both obtained by imagining B shrunk to a single point.

(7.6) Let R denote the continuum obtained by taking a rectangle with bases B_1 and B_2 and altitudes A_1 and A_2 and adding in a sequence of disjoint rectangles, with their interiors, having bases on B_1 and B_2 and converging to A_1. It is readily seen that in this case, the spaces C_2, D_2, C_4, D_4 are single points while each of the spaces D_1, D_3, C_1, C_3 is homeomorphic with the curve obtained by adding the point $(0, 0)$ to that portion of the curve $y^2 = x^2 \sin^2 1/x$ for which $0 < x \leq 1/\pi$.

(7.7) Let us construct a continuum Q as follows. On a diameter U of a circle X choose a non-dense perfect set K containing the end-points of U. Let the segments of $U - K$ be ordered $S_1, S_2, \cdots$ and for each i, let V_i be a circle, together with its interior, having S_i for a diameter. Then let $Q = X + U + \Sigma V_i$. It is readily seen that for this continuum Q, the spaces C_1 and C_2 are θ-curves (i. e., the sum of three arcs having just their end-points in common), D_1 and D_3 are homeomorphic with a lemniscate, while C_2, D_2, C_4 and D_4 reduce to single points.

8. *Totally imperfect and punctiform connected sets. Other ways of defining the sets $C_3(p)$ and $D_3(p)$.*

(8.1) *If the subcontinuum N of a continuum M contains uncountably many points of L, the set of local separating points of M, then every connected subset P of N which is dense in N contains a perfect subset of L.*

For since * $N \cdot L$ is a $G_{\delta\sigma}$, $N \cdot L$ contains † a perfect set K; and since

* See my paper in *Transactions of the American Mathematical Society*, Vol. 32 (1930), p. 180.

† See Hausdorff, *Mengenlehre* (1927), p. 180.

$(\bar{P}-P)\cdot L = X_0$ and $\bar{P}=N$, it follows that $K\cdot P$ contains a perfect set.

(8. 2) *If every subcontinuum of a continuum M contains uncountably many of the local separating points L of M, then every connected subset of M contains a perfect subset of L.*

For if P is any connected set in M, we have only to set $\bar{P}=N$ and apply (8. 1).

(8. 3) *In order that a continuum M contain a totally imperfect connected set it is necessary and sufficient that the local separating points of some subcontinuum of M be countable.*

The necessity follows from (8. 2). The sufficiency has been proved elsewhere.*

COROLLARY. *If every connected subset of M contains a perfect set, then every connected subset P of M contains a perfect set of local separating points of $\bar{P}$.*

(8. 4) *For each $p\epsilon M$, $C_3(p)$ is identically the sum of all totally imperfect connected subsets of M containing p. Likewise if M is hereditarily locally connected, $D_3(p)$ is the sum of all punctiform connected subsets of M containing p.*

For by definition the local separating points of $C_3(p)$ are countable. Hence, by the above, $C_3(p)$ is the closure of a totally imperfect connected set P. Thus for any $x\epsilon C_3(p)$, $P+p+x$ is connected and totally imperfect. On the other hand if P is any connected and totally imperfect subset of M containing p, then since, by (8. 1), $L_3(\bar{P}) = X_0$, we have $P\subset C_3(p)$.

A similar argument establishes the second part of the theorem, using the author's † result that a hereditarily locally connected continuum contains a punctiform connected set if and only if the set of local separating points of some subcontinuum is punctiform.

COROLLARY. *The decomposition space C_3 contains no totally imperfect connected set and, in case M is hereditarily locally connected, D_3 contains no punctiform connected set. Conversely if M is hereditarily locally connected and (a) contains no totally imperfect connected set or (b) contains no punctiform connected set then (a) for each $p\epsilon M$, $C_3(p)=p$* and $M\equiv C_3$, *(b) $D_3(p)=p$ and $D_3\equiv M$.*

---

* See my paper, *American Journal of Mathematics*, Vol. 55 (1933), p. 148.

† *Loc. cit.*, p. 150.

(8.5) *In order that the local separating points of a plane, locally connected and cyclicly connected continuum M be* $\left\{\begin{array}{l}\textit{countable}\\ \textit{punctiform}\end{array}\right\}$ *it is necessary and sufficient that the intersection of the boundaries of every pair of complementary domains of M be* $\left\{\begin{array}{l}\textit{countable}\\ \textit{punctiform}\end{array}\right\}$.

The necessity of both conditions follows from the fact that any point common to the boundaries of two complementary domains of M is a local separating point of M. The sufficiency of the first condition results from the fact that each local separating point of M is common to the boundaries of some pair of complementary domains of M and that there are only a countable number of possible such pairs. The latter condition is sufficient because any continuum of local separating points of M contains a free arc and any free arc of M must be on the boundary of two complementary domains of M.

COROLLARY. *If M is in the plane and is hereditarily locally connected and cyclicly connected, then for each $p \epsilon M$* $\left\{\begin{array}{l}C_3(p)\\ D_3(p)\end{array}\right\}$ *is the maximal subcontinuum N of M containing p and such that the intersection of the boundaries of each pair of the complementary domains of N is* $\left\{\begin{array}{l}\textit{countable}\\ \textit{punctiform}\end{array}\right\}$.

9. *Extensions and additional applications.* Comparing the two sets of conditions given in § 1 we see that they differ essentially only in that one set contains X_0 where the other contains P_σ. This fact suggests the possibility of finding a single set of conditions which could be stated in terms of an independent property P (instead of either X_0 or P_σ) which could be subjected to certain auxiliary restrictions which would be satisfied by both X_0 and P_σ and would be sufficient to prove the theorems in §§ 1-4. Although our original hopes in this direction have not been realized, yet it is possible to obtain some results using an undefined property P.

(9.1) Let P be a property such that any single point has property P and which is countably additive, i. e., the sum of any countable number of sets having property P has property P, and let $L(C)$ be a set-function defined for subcontinua C of M and such that *

(i) $L(C) \subset C$

(ii) $L(\bar{E}) \cdot (\bar{E} - E) = P$ (i. e., has property P, where E is any connected set)

(iii) $L(C') = C' \cdot L(C)$, where C' is a subcontinuum of C.

* Here of course we have (iii) → (i).

Then by the same method as used in § 2 we can show that for each $p \epsilon M$ there exists a maximal subcontinuum $G(p)$ of M containing p and such that $L[G(p)]$ has property P. Likewise $q \epsilon G(p)$ implies $G(q) \equiv G(p)$, so that the sets $G(p)$ are disjoint; and furthermore the decomposition of M into sets $G(p)$ is upper semi-continuous.

(9.2) As an application of this, let P be the property of being a P_σ and let $L(C)$ be defined as the set of local separating points of M belonging to C. Then (i), (ii), (iii) are satisfied, and for each $p \epsilon M$, $G(p)$ is the maximal subcontinuum of M containing p and such that $G(p) \cdot L(M)$ is a P_σ. Hence $G(p) = D_1(p)$ provided $L(M)$ is an F_σ. As a second application, we could take P to be the property of being countable and, using the function $L_1(C)$ of § 5, obtain $G(p) = C_1(p)$.

(9.3) *If N is any subcontinuum of M having no cut point, there exists a maximal subcontinuum $K(N)$ of M containing N and having no cut point.*

Proof. Let $K(N)$ be the sum of all subcontinua H of M having no cut point and such that $H \cdot N$ contains at least two points. Then clearly $K(N)$ is connected; and since no point of $\overline{K(N)} - K(N)$ could cut $\overline{K(N)}$ we have $K(N) = \overline{K(N)}$, so that $K(N)$ is a continuum. Now if $K(N)$ had a cut point p, p would separate some point x of $K(N) - N$ from every point of $N - N \cdot p$; but x lies together with some point y of $N - N \cdot p$ in a subcontinuum H of $K(N)$ which has no cut point, which clearly is impossible. Thus $K(N)$ has no cut point; and since $K(N)$, by definition, contains all subcontinua of M which contain N and have no cut point, clearly it is the maximal subcontinuum having this property.

(9.4) Now by the method of finding the set $K(N)$ it follows at once that if N_1 and N_2 are two subcontinua of M having no cut points, then either $K(N_1) \equiv K(N_2)$, $K(N_1) \cdot K(N_2) = 0$, or $K(N_1) \cdot K(N_2)$ reduces to a single point, and all three cases are possible. Likewise it is readily shown that any set $K(N)$ contains at most a countable number of points x such that x belongs to more than one set $K(N)$, and that $a, b \epsilon K(N)$ implies that every irreducible subcontinuum between a and b is contained in $K(N)$.

10. *Decompositions of cyclic elements.* Let us suppose that M is locally connected and consider the relations between the sets $C_i(p)$ and $D_i(p)$ $(i = 1, 2, 3, 4)$ in M and the non-degenerate cyclic elements E of M. In the first place it is apparent that since no E has a cut point of itself nor does it contain but a countable number of cut points of M, each E is contained wholly in $C_2(p)$, $C_4(p)$, $D_2(p)$, and $D_4(p)$, where p is any point of E. However,

when we consider the decompositions into sets $C_i(p)$ and $D_i(p)$, (i odd), given by local separating points we see by examining examples such as (7.7) that the cyclic elements of M and the sets $C_i(p)$ and $D_i(p)$ are independent in the sense that either may be contained wholly or only partially in the other. This suggests the possibility of obtaining "finer" decompositions of M by alternating the cyclic element decomposition of M with the decompositions into set $C_i(p)$ and $D_i(p)$. For example, we may either first decompose M into sets $C_1(p)$ and then consider the cyclic elements of the hyperspace C_1, or first decompose M into cyclic elements and then decompose each of these elements into sets $C_1(p)$. Then either of these steps may be repeated; however, in any case the decomposition stops (i.e., each $C_1(p)$ reduces to p) as soon as we have performed the $C_1(p)$ decomposition and then the cyclic element decomposition in this order.

If we make use of the facts (i) that if N is any subcontinuum of M, $N \cdot E$ is either vacuous or connected, (ii) that any local separating point of $N \cdot E$ is a local separating point of N, and (iii) that any point of $N \cdot E$ which is a local separating point of N either is a cut point of M or a local separating point of $N \cdot E$, we can prove immediately

(10.1) *If $C_i^e(p)$ and $C_i^m(p)$ [also $D_i^e(p)$, $D_i^m(p)$], denote decompositions of E and M respectively, then for each non-degenerate cyclic element E of M and each $p \epsilon E$, we have*

$$C_i^e(p) \equiv E \cdot C_i^m(p),$$
$$D_i^e(p) \equiv E \cdot D_i^m(p) \qquad (i = 1, 2, 3, 4),$$

where for $i = 3$, in the D decomposition it is supposed that M is hereditarily locally connected.

In conclusion the author wishes to direct attention to the possibility of discovering a group of properties of continua which would be extendible from the sets $C_i(p)$ and $D_i(p)$ to the whole continuum in the same sense that the cyclicly extendible properties of locally connected continua extend from the cyclic elements to the whole continuum. Also it would be desirable to develop in greater detail the properties of the types of continua C_i and D_i which we obtained as decomposition spaces in § 5.

THE JOHNS HOPKINS UNIVERSITY.

MINIMAL SURFACES IN EUCLIDEAN *N*-SPACE.

By E. F. Beckenbach.*

1. *Introduction.* Let the rectangular coördinates of a surface in euclidean n-space be given by

$$x_r = x_r(U, V), \qquad (r = 1, 2, \cdots, n).$$

If the minimal curves are parametric, so that

$$E = G = 0,$$

then a necessary and sufficient condition that the surface be minimal is that

$$\partial^2 x_r / \partial U \, \partial V = 0.$$

This gives

$$x_r = U_r(U) + V_r(V); \tag{1}$$

and, since the minimal curves are parametric,

$$\sum_{r=1}^{n} U'^2_r = 0, \qquad \sum_{r=1}^{n} V'^2_r = 0, \tag{2}$$

the primes denoting differentiation with respect to the respective arguments U and V.

For $n = 3$, the Enneper-Weierstrass equations,

$$x_1 = (1/2) \int (1 - u^2) F(u) \, du + (1/2) \int (1 - v^2) \Phi(v) \, dv,$$

$$x_2 = (i/2) \int (1 + u^2) F(u) \, du - (i/2) \int (1 + v^2) \Phi(v) \, dv,$$

$$x_3 = \int u F(u) \, du + \int v \Phi(v) \, dv,$$

are obtained from (1) and (2) by writing

$$\frac{dU_1 + i dU_2}{-dU_3} = \frac{dU_3}{dU_1 - i dU_2} = u, \quad \frac{-dV_3}{dV_1 + i dV_2} = \frac{dV_1 - i dV_2}{dV_3} = v.$$

Eisenhart † obtained analogous formulae for $n = 4$ by noting that in this case (2) may be written in the form

---

* National Research Fellow.

† *Annals of Mathematics*, Ser. 2, Vol. 13 (1911), pp. 17-35; *American Journal of Mathematics*, Vol. 34 (1912), pp. 215-236.

$$\frac{dU_1 + idU_2}{dU_3 + idU_4} = -\frac{dU_3 - idU_4}{dU_1 - idU_2} = u,$$

with similar equations for v.

The purpose of the present paper is to give analogous formulae for a general n and to point out the fact that several results which follow from the Enneper-Weierstrass equations follow also from these generalized equations.

2. *Normal parametric representation of minimal curves.* Let the rectangular coördinates of a curve be given by

$$x_r = U_r(U), \qquad (r = 1, 2, \cdots, n).$$

A minimal curve, or curve of zero length, is a curve for which

$$\sum_{r=1}^{n} U'^2_r = 0. \tag{3}$$

Equation (3) can be written as

$$\frac{dU_1 + idU_2}{-(\sum_{r=3}^{n} dU_r^2)^{1/2}} = \frac{(\sum_{r=3}^{n} dU_r^2)^{1/2}}{dU_1 - idU_2} = u^{1/2} = [F_1(u)]^{1/2}. \tag{4}$$

We neglect for the present the possibility that u is either constant or indeterminate, and call u the normal parameter of the curve. And we call the functions $F_r(u)$, $r = 2, 3, \cdots, n-1$, which we now shall determine, the normal functions of the curve.

The above definition of u yields

$$dU_1 : dU_2 : (\sum_{r=3}^{n} dU_r^2)^{1/2} = (1/2)(1 - F_1) : (i/2)(1 + F_1) : F_1^{1/2}, \tag{5}$$

so that, neglecting constants of integration which can be removed by a translation, we have

$$\begin{aligned} U_1 &= (1/2) \int (1 - F_1) F_2 \, du, \\ U_2 &= (i/2) \int (1 + F_1) F_2 \, du, \\ (\sum_{r=3}^{n} U'^2_r)^{1/2} &= F_1^{1/2} F_2, \end{aligned} \tag{6}$$

where $F_2(u)$ is the function of proportionality defined by

$$F_2(u) = \frac{2U'_1}{1 - F_1} = \frac{-2iU'_2}{1 + F_1} = \left(\frac{\sum_{r=3}^{n} U'^2_r}{F_1} \right)^{1/2},$$

the prime now denoting differentiation with respect to u.

We start now with the equation

$$\sum_{r=3}^{n} U'^2_r - F_1 F_2^2 = 0,$$

and proceed exactly as before to determine U_3 and U_4 in terms of F_3 and F_4, and so on. In general, we start with

$$\sum_{r=2s-1}^{n} U'^2_r - \sum_{r=1}^{s-1} F_{2r-1} F^2_{2r} = 0,$$

and define

$$F_{2s-1}(u) = \frac{(U'_{2s-1} + iU'_{2s})^2}{\sum_{r=2s+1}^{n} U'^2_r - \sum_{r=1}^{s-1} F_{2r-1} F^2_{2r}} = \frac{\sum_{r=2s+1}^{n} U'^2_r - \sum_{r=1}^{s-1} F_{2r-1} F^2_{2r}}{(U'_{2s-1} - iU'_{2s})^2},$$

$$(7) \qquad F_{2s}(u) = \frac{2U'_{2s-1}}{1 - F_{2s-1}} = \frac{-2iU'_{2s}}{1 + F_{2s-1}} = \left(\frac{\sum_{r=2s+1}^{n} U'^2_r - \sum_{r=1}^{s-1} F_{2r-1} F^2_{2r}}{F_{2s-1}} \right)^{1/2},$$

so that

$$(8) \quad U_{2s-1} = (1/2) \int (1 - F_{2s-1}) F_{2s} du, \quad U_{2s} = (i/2) \int (1 + F_{2s-1}) F_{2s} du.$$

Finally, if n is even, $n = 2m$, we have

$$F_{2m} = \left(\frac{-\sum_{r=1}^{m-1} F_{2r-1} F^2_{2r}}{F_{2m-1}} \right)^{1/2},$$

whence

$$U_{2m-1} = (1/2) \int (1 - F_{2m-1}) \left(\frac{-\sum_{r=1}^{m-1} F_{2r-1} F^2_{2r}}{F_{2m-1}} \right)^{1/2} du,$$

$$(9) \qquad U_{2m} = (i/2) \int (1 + F_{2m-1}) \left(\frac{-\sum_{r=1}^{m-1} F_{2r-1} F^2_{2r}}{F_{2m-1}} \right)^{1/2} du;$$

while if n is odd, $n = 2m + 1$, we have

$$(10) \qquad U_{2m+1} = \int \left(\sum_{r=1}^{m} F_{2r-1} F^2_{2r} \right)^{1/2} du.$$

It is to be noted that we do not have two alternatives in selecting the roots appearing in (9) and (10), since we must choose those roots which yield the given U_r involved.

Whether n is even or odd, then, we have expressed the n functions U_r in terms of the unique parameter u and the $n - 2$ unique functions $F_2, \cdots,$

F_{n-1}. Conversely, any $n-2$ analytic functions put into these equations determine a minimal curve in n dimensional euclidean space.

Our parametric representation serves also for the expression of the coördinates of a space curve in $n-1$ dimensions, for if we let s represent the arc length and set

$$x_n = is,$$

then (3) becomes

$$ds^2 = \sum_{r=1}^{n-1} dx_r^2.$$

3. *The exceptional case.* If u is constant, $u = c$, we see by (6) that

$$U_1 = (1/2)(1-c)g(U), \qquad U_2 = (i/2)(1+c)g(U),$$

where $g(U)$ is the integral of the function of proportionality in (5). The projection of the minimal curve on the (x_1, x_2)-plane is therefore a straight line.

If u is indeterminate, then x_1 and x_2 are both constants or

$$x_2 = \pm ix_1 + k$$

where k is a constant, so that the projection on the (x_1, x_2)-plane is either a point or a straight line.

Conversely, if the projection of the curve on this plane is a point or a straight line, u is either constant or indeterminate. For if x_1 and x_2 are both constant, u is indeterminate; and if

$$ax_1 + bx_2 + c = 0,$$

where not both a and b are zero, then either

$$u = (a + ib)/(a - ib) = c,$$

or, if $a \pm ib = 0$, u is indeterminate.

If the projection on the (x_1, x_2)-plane is a point or a straight line, we define

$$(dU_1 + idU_3)^2 / \sum_{r \neq 1,3} dU_r^2 = u,$$

provided this quantity is neither constant nor indeterminate. We call u the normal parameter and proceed to determine the normal functions for the sequence

$$U_1, U_3, U_2, U_4, U_5, \cdots, U_n$$

just as we would do otherwise for

$$U_1, U_2, \cdots, U_n.$$

In general, if x_p is the coördinate of lowest rank for which there exists at least one other coördinate x_t such that

$$(11) \qquad (dU_p + idU_t)^2 / \sum_{r \neq p,t} dU_r^2$$

is neither constant nor indeterminate, and if x_s is the coördinate of lowest rank among all such x_t, we determine the normal parameter and normal functions for the sequence

$$U_p, U_s, U_1, U_2, \cdots, U_n,$$

and define these to be the normal parameter and normal functions for the minimal curve.

If for all U_p and U_t, (11) is either constant or indeterminate, our normal parametric representation is impossible. In this case, each coördinate of the curve is a linear function of each other coördinate, excepting that some might be identically constant, and the minimal curve is a straight line or a point. Conversely, if the minimal curve is a straight line or a point, then for all U_p and U_t, (11) is either constant or indeterminate and the normal parametric representation is impossible.

Every minimal straight line,

$$x_r = \alpha_r U + \beta_r,$$

lies on the minimal cone, or sphere of zero radius,

$$(12) \qquad \sum_{r=1}^{n} (x_r - \beta_r)^2 = 0,$$

so that if a minimal curve cannot be given in normal parametric representation, it lies on (12).

4. *Reflections in the coördinate hyperplanes.* If we reflect a minimal curve in the hyperplane $x_r = 0$, we obtain a minimal curve the coördinates of which are the same functions as those of the original curve except the r-th, which differs from the original in sign only. We shall have use in the next section for the relations between the normal parameter and normal functions of the original curve, which we denote by u and $F_r(u)$, and those of the reflection, which we denote by $\mathfrak{U}$ and $\mathfrak{F}_r(\mathfrak{U})$.

If we reflect in $x_1 = 0$, we obtain

$$\mathfrak{U} = 1/u, \qquad \mathfrak{F}_2(\mathfrak{U}) = uF_2(u)\,du/d\mathfrak{U},$$
$$\mathfrak{F}_{2s-1}(\mathfrak{U}) = F_{2s-1}(u), \qquad \mathfrak{F}_{2s}(\mathfrak{U}) = F_{2s}(u)\,du/d\mathfrak{U}, \qquad s > 1;$$

if we reflect in $x_2 = 0$, we have

$$\mathfrak{U} = 1/u, \qquad \mathfrak{F}_2(\mathfrak{U}) = -uF_2(u)\,du/d\mathfrak{U},$$
$$\mathfrak{F}_{2s-1}(\mathfrak{U}) = F_{2s-1}(u), \qquad \mathfrak{F}_{2s}(\mathfrak{U}) = F_{2s}(u)\,du/d\mathfrak{U}, \qquad s > 1.$$

Reflecting in $x_{2r-1} = 0$, $r > 1$, we get $\mathfrak{U} = u$ and

$$\mathfrak{F}_s(\mathfrak{U}) = F_s(u), \qquad s \neq 2r - 1, \ s \neq 2r,$$
$$\mathfrak{F}_{2r-1} = 1/F_{2r-1}, \qquad \mathfrak{F}_{2r} = F_{2r-1}F_{2r},$$

while reflecting in $x_{2r} = 0$, $r > 1$, we have $\mathfrak{U} = u$ and

$$\mathfrak{F}_s(\mathfrak{U}) = F_s(u), \qquad s \neq 2r - 1, \ s \neq 2r,$$
$$\mathfrak{F}_{2r-1} = 1/F_{2r-1}, \qquad \mathfrak{F}_{2r} = -F_{2r-1}F_{2r}.$$

If $n = 2m$, and we reflect in the hyperplane $x_{2m-1} = 0$, we obtain

$$\mathfrak{F}_s = F_s, \quad s < 2m - 1, \qquad \mathfrak{F}_{2m-1} = 1/F_{2m-1},$$
$$\left(\frac{-\sum_{r=1}^{m-1} \mathfrak{F}_{2r-1}\mathfrak{F}^2_{2r}}{\mathfrak{F}_{2m-1}}\right)^{1/2} = F_{2m-1}\left(\frac{-\sum_{r=1}^{m-1} F_{2r-1}F^2_{2r}}{F_{2m-1}}\right)^{1/2},$$

while if we reflect in $x_{2m} = 0$, we obtain

$$\mathfrak{F}_s = F_s, \quad s < 2m - 1, \qquad \mathfrak{F}_{2m-1} = 1/F_{2m-1},$$
$$\left(\frac{-\sum_{r=1}^{m-1} \mathfrak{F}_{2r-1}\mathfrak{F}^2_{2r}}{\mathfrak{F}_{2m-1}}\right)^{1/2} = -F_{2m-1}\left(\frac{-\sum_{r=1}^{m-1} F_{2r-1}F^2_{2r}}{F_{2m-1}}\right)^{1/2}.$$

Finally, if $n = 2m + 1$, and we wish to reflect in the hyperplane $x_{2m+1} = 0$, we have but to choose the negative of the square root appearing in (10).

In solving for the F_r in any of the above reflections, we note that the F_r are the same functions of the $\mathfrak{F}$'s as the $\mathfrak{F}_r$ are of the F's.

We can reflect in as many of the coördinate hyperplanes as we wish, making the reflections one at a time; the equations of the transformation result from the succession of the separate sets of equations. For example, if we reflect in both $x_{2r-1} = 0$ and $x_{2r} = 0$, we have

$$\mathfrak{F}_s = F_s, \quad s \neq 2r, \qquad \mathfrak{F}_{2r} = -F_{2r}.$$

We note that as a result of these reflection formulae we can, by a suitable change of the normal functions, express the coördinates of a minimal curve by equations of the form (6), (8), (9), (10), except that such of the integrals as we please are multiplied by minus one. We shall use this fact in the next section.

5. *Normal parametric representation of minimal surfaces.* Since the

functions $U_r(U)$ and $V_r(V)$ in (1) satisfy (2), these two sets of functions each can be given in normal representation. Let the normal functions be respectively $F_r(u)$ and $\Psi_r(t)$. According to section 4, we can replace the functions $\Psi_r(t)$ by the functions $\Phi_r(v)$ so that the functions (1) representing a minimal surface in n dimensions can be written in the form:

$$(13)\quad \begin{aligned} x_{2s-1} &= (1/2)\int (1-F_{2s-1})F_{2s}\,du + (1/2)\int (1-\Phi_{2s-1})\Phi_{2s}\,dv, \\ x_{2s} &= (i/2)\int (1+F_{2s-1})F_{2s}\,du - (i/2)\int (1+\Phi_{2s-1})\Phi_{2s}\,dv; \end{aligned}$$

if $n = 2m$,

(14)

$$x_{2m-1} = (1/2)\int (1-F_{2m-1})\left[\frac{-\sum_{r=1}^{m-1} F_{2r-1}F^2_{2r}}{F_{2m-1}}\right]^{1/2} du + (1/2)\int (1-\Phi_{2m-1})\left[\frac{-\sum_{r=1}^{m-1} \Phi_{2r-1}\Phi^2_{2r}}{\Phi_{2m-1}}\right]^{1/2} dv,$$

$$x_{2m} = (i/2)\int (1+F_{2m-1})\left[\frac{-\sum_{r=1}^{m-1} F_{2r-1}F^2_{2r}}{F_{2m-1}}\right]^{1/2} du - (i/2)\int (1+\Phi_{2m-1})\left[\frac{-\sum_{r=1}^{m-1} \Phi_{2r-1}\Phi^2_{2r}}{\Phi_{2m-1}}\right]^{1/2} dv;$$

if $n = 2m+1$,

$$(15)\quad x_{2m+1} = \int \Big(\sum_{r=1}^{m} F_{2r-1}F^2_{2r}\Big)^{1/2} du + \int \Big(\sum_{r=1}^{m} \Phi_{2r-1}\Phi^2_{2r}\Big)^{1/2} dv.$$

Here the F_r are given by (4) and (7), while, according to section 4, the Φ_r are given by

$$\Phi_1(v) = v = \frac{\sum_{r=3}^{n} V'^2_r}{(V'_1 + iV'_2)^2} = \frac{(V'_1 - iV'_2)^2}{\sum_{r=3}^{n} V'^2_r},$$

$$\Phi_{2s-1}(v) = \frac{\sum_{r=2s+1}^{n} V'^2_r - \sum_{r=1}^{s-1} \Phi_{2r-1}\Phi^2_{2r}}{(V'_{2s-1} + iV'_{2s})^2} = \frac{(V'_{2s-1} - iV'_{2s})^2}{\sum_{r=2s+1}^{n} V'^2_r - \sum_{r=1}^{s-1} \Phi_{2r-1}\Phi^2_{2r}},$$

$$\Phi_{2s}(v) = \frac{2V'_{2s-1}}{1-\Phi_{2s-1}} = \frac{2iV'_{2s}}{1+\Phi_{2s-1}} = \left[\frac{\sum_{r=2s+1}^{n} V'^2_r - \sum_{r=1}^{s-1} \Phi_{2r-1}\Phi^2_{2r}}{\Phi_{2s-1}}\right]^{1/2},$$

the prime denoting differentiation with respect to v.

If and only if the coördinates of the minimal surface are given by these equations (13), (14), (15), we say that u and v are the normal parameters, and the F_r and Φ_r are the normal functions, of the surface.

According to the discussion of section 3, if it is impossible thus to choose one of the normal parameters, say u, then the curves, $v = \text{constant}$, on the surface are parallel straight lines: the surface is a cylinder. If neither parameter can be determined, the cylinder is a plane.

In terms of these functions, the fundamental quantities of the first order are

$$E = G = 0$$

and

$$F = (1/2)\sum_{r=1}^{m-1} F_{2r}\Phi_{2r}(1 + F_{2r-1}\Phi_{2r-1})$$
$$+ (1/2)\left[\frac{(-\sum_{r=1}^{m-1} F_{2r-1}F^2_{2r})(-\sum_{r=1}^{m-1}\Phi_{2r-1}\Phi^2_{2r})}{F_{2m-1}\Phi_{2m-1}}\right]^{1/2}[1 + F_{2m-1}\Phi_{2m-1}],$$

or

$$F = (1/2)\sum_{r=1}^{m} F_{2r}\Phi_{2r}(1 + F_{2r-1}\Phi_{2r-1}) + [(\sum_{r=1}^{m} F_{2r-1}F^2_{2r})(\sum_{r=1}^{m}\Phi_{2r-1}\Phi^2_{2r})]^{1/2},$$

according as $n = 2m$ or $n = 2m + 1$.

6. *Real minimal surfaces.* The above particular form (13), (14), (15) of the expression of the coördinates of a minimal surface was chosen, as regards plus and minus signs, because if $F_r(u)$ and $\Phi_r(v)$, $r = 1, 2, \cdots, n - 1$, are conjugate imaginaries, and if, for $n = 2m$,

$$(16) \qquad \left(\frac{-\sum_{r=1}^{m-1} F_{2r-1}F^2_{2r}}{F_{2m-1}}\right)^{1/2} \quad \text{and} \quad \left(\frac{-\sum_{r=1}^{m-1}\Phi_{2r-1}\Phi^2_{2r}}{\Phi_{2m-1}}\right)^{1/2}$$

are conjugates, or if, for $n = 2m + 1$,

$$(17) \qquad (\sum_{r=1}^{m} F_{2r-1}F^2_{2r})^{1/2} \quad \text{and} \quad (\sum_{r=1}^{m}\Phi_{2r-1}\Phi^2_{2r})^{1/2}$$

are conjugates, then the surface is real. This follows at once from the fact that to each element of the integral relative to u there corresponds, in each of the n equations, an element in the integral relative to v which is its conjugate imaginary. In this case we may write

$$x_{2s-1} = R \int (1 - F_{2s-1}) F_{2s}\, du,$$

$$x_{2s} = R \int i(1 + F_{2s-1}) F_{2s}\, du;$$

if $n = 2m$,

$$x_{2m-1} = R \int (1 - F_{2m-1}) \left[\frac{-\sum_{r=1}^{m-1} F_{2r-1} F^2_{2r}}{F_{2m-1}} \right]^{1/2} du,$$

$$x_{2m} = R \int i(1 + F_{2m-1}) \left[\frac{-\sum_{r=1}^{m-1} F_{2r-1} F^2_{2r}}{F_{2m-1}} \right]^{1/2} du;$$

if $n = 2m + 1$,

$$x_{2m+1} = R \int 2 \left(\sum_{r=1}^{m} F_{2r-1} F^2_{2r} \right)^{1/2} du,$$

where R designates "the real part of."

We shall show now that the conditions that F_r and Φ_r, and the functions in (16) or (17), be conjugate imaginaries are also necessary in order that the surface be real. For a real minimal surface, $H \equiv (EG - F^2)^{1/2}$ can vanish only at isolated points. In a small region about any other point, then,

$$\partial(x_a, x_b)/\partial(U, V) \neq 0$$

for some a, b; consequently, U and V are functions of x_a and x_b in this region. Let

$$dx_t = p_t dx_a + q_t dx_b. \tag{18}$$

Along the minimal curves, which we are taking to be parametric, we have

$$\sum_{r=1}^{n} dx_r^2 = 0. \tag{19}$$

Equations (18) and (19) yield

$$\begin{aligned} dx_a : dx_b : dx_t = \{ &- \sum_{r=1}^{n} p_r q_r \pm i [\sum_{r<s} (p_r q_s - p_s q_r)^2]^{1/2} \} \\ &: \sum_{r=1}^{n} p_r^2 : \{ - p_t \sum_{r=1}^{n} p_r q_r + q_t \sum_{r=1}^{n} p_r^2 \pm i p_t [\sum_{r<s} (p_r q_s - p_s q_r)^2]^{1/2} \}. \end{aligned} \tag{20}$$

Since for real surfaces the p_t and q_t are real, the corresponding terms of the two systems of ratios in the right-hand member of (20) are conjugate imaginaries. By their equations of definition, then, the F_r and Φ_r are conjugate imaginaries. The quantities in (16) or (17) therefore must be conjugate imaginaries or negative conjugate imaginaries; but in the latter case,

for $n = 2m$, x_{2m-1} and x_{2m} would be pure imaginaries, or, for $n = 2m + 1$, x_{2m+1} would be a pure imaginary, contrary to hypothesis.

7. *Associate minimal surfaces.* The surfaces S_α, whose coördinates are given by

$$x_{r,\alpha} = e^{i\alpha}U_r(U) + e^{-i\alpha}V_r(V),$$

where $U_r(U)$ and $V_r(V)$ are the functions in (1), form a one parameter family of minimal surfaces applicable each to each, and called *associate minimal surfaces.* The linear element of each of them is given by

$$ds^2 = \sum_{r=1}^{n} dx_r^2 = 2\sum_{r=1}^{n} dU_r dV_r.$$

The normal functions defining S_α are

$$F_{2s-1}, \quad e^{i\alpha}F_{2s}, \quad \Phi_{2s-1}, \quad e^{-i\alpha}\Phi_{2s},$$

where

$$F_{2s-1}, \quad F_{2s}, \quad \Phi_{2s-1}, \quad \Phi_{2s}$$

are the normal functions of S_0 defined by (1).

The Jacobians $J_{rs} = \partial(x_r, x_s)/\partial(u, v)$ are the same for S_0 and S_α, and consequently so are the direction-cosines $P_{rs} = J_{rs}/H$ of the tangent planes. The tangent planes at corresponding points of a family of associate minimal surfaces are therefore parallel.

The surface $S_{\pi/2}$, whose coördinates we designate by y_r, is called the *adjoint* of S_0. We have

$$x_{r,\alpha} = x_r \cos\alpha + y_r \sin\alpha.$$

Since the tangent planes to S_0 and $S_{\pi/2}$ are parallel at corresponding points, we have

$$(21) \qquad \begin{Vmatrix} dy_1, & dy_2, & \cdots, & dy_n \\ \dfrac{\partial x_1}{\partial u}, & \dfrac{\partial x_2}{\partial u}, & \cdots, & \dfrac{\partial x_n}{\partial u} \\ \dfrac{\partial x_1}{\partial v}, & \dfrac{\partial x_2}{\partial v}, & \cdots, & \dfrac{\partial x_n}{\partial v} \end{Vmatrix} = 0.$$

We have also

$$(22) \qquad \sum_{r=1}^{n} dx_r dy_r = 0,$$

so that corresponding curves on a minimal surface and on its adjoint are perpendicular to one another at corresponding points.

By means of (13), (14), (15) and (21), (22), we obtain

$$\frac{dy_1}{\sum_{r=1}^{n} P_{1r}dx_r} = \frac{dy_2}{\sum_{r=1}^{n} P_{2r}dx_r} = \cdots = \frac{dy_n}{\sum_{r=1}^{n} P_{nr}dx_r} = -1,$$

where we have taken $H = iF$. We have therefore

$$\begin{aligned} x_s - iy_s &= x_s + i \int \sum_{r=1}^{n} P_{sr}dx_r = 2U_s, \\ x_s + iy_s &= x_s - i \int \sum_{r=1}^{n} P_{sr}dx_r = 2V_s. \end{aligned} \tag{23}$$

These formulae (23) are analogous to the formulae of Schwarz for minimal surfaces in ordinary space. By means of them we verify readily that the coördinates of a minimal surface, passing through a curve whose coördinates are given by the analytic functions $x_s = x_s(t)$ and admitting at each point of the curve a tangent plane whose direction-cosines are given by the analytic functions $P_{sr}(t)$, are given by

$$X_s = (1/2)[x_s(u) + x_s(v)] + (i/2) \sum_{r=1}^{n} \int_v^u P_{sr}dx_r.$$

From this last, we obtain the following two results.

If a straight line lies on a minimal surface, it is an axis of symmetry of the surface.

If a minimal surface cuts a hyperplane, say the $(x_1, x_2, \cdots, x_k)$ hyperplane, normally, it is symmetric with respect to the hyperplane.

THE RICE INSTITUTE,
HOUSTON, TEXAS.

ON GENERALIZED MANIFOLDS.

By S. Lefschetz.

The object of the present paper is to extend to a larger class of spaces certain results recently obtained for topological manifolds.† The extension consists in replacing the requirement that every point possess a combinatorial cell for neighborhood by certain weaker conditions on the chains through the point. Roughly speaking they amount to demanding that locally any p-chain be deformable (in a certain very general sense) into one which does not meet any assigned q-space ($= q$ dimensional space), where $p + q < n$, the dimension of the manifold. This extension is made in Part III of the present paper. In Part I we take up again, partly as a preparation to the second Part, the homology theory of metric spaces from the standpoint initiated in our Colloquium Lectures *Topology,* Ch. VII. The notation and terminology are as in our book.‡

§ 1. The Approximating Complexes of a Metric Space.

1. The homology properties of a compact metric space are intimately related to the homology properties of certain subchains of an infinite complex, the fundamental complex of the space (*Topology,* Ch. VII), or to certain sequences of chains of approximating complexes (Alexandroff). We shall first show how these may be selected in a certain convenient way for the sequel.

Let for the present $\mathcal{R}$ be a compact metric n-space and let $U, V, W,$ denote generically its open sets, and $F(U)$, $F(V)$, $F(W)$, their boundaries.

We shall repeatedly consider various aggregates of subsets, $\Sigma = \{A^{\alpha}\}$, of $\mathcal{R}$. The mesh of Σ is max diam A^{α}. If the set of A's covers $\mathcal{R}$ we call Σ a *covering,* an *ϵ-covering* if its mesh $< \epsilon$. Of particular importance are the finite coverings by open sets ($=$ f. c. o. s.).

Each set A^{α} of the aggregate Σ may be considered as an abstract point,

† S. Lefschetz and W. W. Flexner, *Proceedings of the National Academy,* Vol. 16 (1930), pp. 530-533; W. W. Flexner, *Annals of Mathematics,* Ser. 2, Vol. 32 (1931), pp. 393-406, 539-548.

‡ A very extensive paper by Cěch on the same general topic was presented simultaneously with the present one to the *Annals of Mathematics* where his paper is now appearing. While there are many contacts between the two, they differ essentially in method and scope. Cěch deals indeed with a much more general type of space, but the restriction to locally compact metric spaces which we have imposed here, has enabled us to proceed much more quickly to the point.

and we may then introduce for each intersection $A^{a_0} \cdots A^{a_p} \neq 0$ an abstract p-simplex $\sigma_p = A^{a_0} \cdots A^{a_p}$. It will be convenient to designate the intersection also by σ_p : $\sigma_p = 0$ signifies then that the sets $A^{a_0}, \cdots, A^{a_p}$ do not intersect.

The aggregate $\{\sigma\}$ has the property that with each σ every face of σ also belongs to the set. Hence $\{\sigma\}$ is a closed simplicial (abstract) complex Φ, the *skeleton* of Σ. If another aggregate $\Sigma' = \{A'^{a}\}$ has for skeleton Φ' a complex whose structure is that of a subcomplex of Φ, we shall briefly say that its skeleton is a subcomplex of Φ. The *dimension* of Φ is the highest integer ν such that there is at least one aggregate of $\nu + 1$ intersecting A's. ν is also called the *order* of Σ. Clearly of course Φ is finite when and only when Σ is finite.

Suppose in particular that $\Sigma = \{U^a\}$ is an ϵ-f. c. o. s. It is called *irreducible* (Alexandroff) when there is no ϵ-f. c. o. s. whose skeleton is a proper subcomplex of Φ. If Σ is reducible there is an ϵ-f. c. o. s. Σ^1 whose skeleton is a proper subcomplex Φ^1 of Φ. If Σ^1 is in turn reducible there is an ϵ-f. c. o. s. Σ^2 whose skeleton is a proper subcomplex Φ^2 of Φ, etc. Since Φ has only a finite number of subcomplexes the process must stop after a finite number of steps. Therefore there exists an irreducible ϵ-f. c. o. s. whatever ϵ. If the order of the initial covering is the least possible for an ϵ-f. c. o. s. it will also be the order of the ultimate irreducible covering.

We recall that as $\epsilon \to 0$ the least order ν tends to an upper limit n or else $\to \infty$. In the first case dim $\mathfrak{R} = n$, in the second case dim $\mathfrak{R} = \infty$.

Let $\Sigma = \{U^a\}$ be a f. c. o. s. whose skeleton Φ is the same as for $\{\bar{U}^a\}$. Then there exists a constant η, the *characteristic constant* of Σ, such that: (a) if a set A on $\mathfrak{R}$ whose diameter $< \eta$ meets a certain number of U's, these U's have a non-vacuous intersection; (b) any point x of $\mathfrak{R}$ is on at least one U such that $d(x, \mathfrak{R} - U) > \eta$. As a consequence of (b) if diam $A < \eta$ then some $U \supset A$.

2. Taking $n = \dim \mathfrak{R}$ finite, let ϵ be so small that the least order of an ϵ-f. c. o. s. is n, and let Σ be an irreducible ϵ-f. c. o. s. There exists another ϵ-f. c. o. s. of order n, $\Sigma' = \{V^a\}$ consisting of as many sets as Σ and such that for every α we have $V^a \Subset U^a$.† Clearly Σ' is an ϵ-f. c. o. s. whose skeleton is Φ or a subcomplex of Φ, and since Σ is irreducible it can only be Φ. Therefore *the order of* Σ *is* n. In other words an irreducible f. c. o. s. whose mesh is sufficiently small is of order n. Observe incidentally that $\{\bar{V}^a\}$ has the same skeleton as $\{V^a\}$.

† Menger, *Dimensionstheorie*, p. 160. We shall use his "strong inclusion" symbol $\Subset$ ($A \Subset B$ means that $\bar{A} \subset B$).

Consider now a sequence $\{\Sigma^i\}$, where $\Sigma^i = \{U^{i\alpha}\}$ is an irreducible ϵ_i-f. c. o. s. such that: (a) $\epsilon_1 = \epsilon$; (b) if η_i is the characteristic constant of Σ^i we have $\epsilon_{i+1} < \frac{1}{2}\eta_i$ and $< \frac{1}{2}\epsilon_i$; (c) $\{\bar{U}^{i\alpha}\}$ has the same skeleton as Σ^i. As a consequence Σ^i is of order n and for every $U^{i+1,\beta}$ there is a $U^{i\alpha} \supset U^{i+1,\beta}$. Let Φ^i be the skeleton of Σ^i; choose for each $U^{i+1,\beta}$ a definite $U^{i\alpha} \supset U^{i+1,\beta}$ and define a transformation t_i of the vertices of Φ^{i+1} into vertices of Φ^i whereby the vertex $U^{i+1,\beta}$ goes into the vertex $U^{i\alpha}$. Let $\sigma_p = U^{i+1,\beta_0} \cdots U^{i+1,\beta_p}$ be a simplex of Φ^{i+1}. As a consequence, if $U^{i\alpha_h} = t_i U^{i+1,\beta_h}$ then $U^{i\alpha_h} \supset U^{i+1,\beta_h}$ and hence $\sigma'_q = U^{i\alpha_0} \cdots U^{i\alpha_p}$ is a simplex of Φ^i. (It may happen that several of the vertices $U^{i\alpha_h}$ coincide, in which case $q < p$). Thus if certain vertices U^{i+1} belong to a σ_p of Φ^{i+1} the transformed vertices $t_i U^{i+1}$ are vertices of a σ_q $(q \leqq p)$ of Φ^i. Consequently t_i may be extended to a simplicial transformation τ_i of Φ^{i+1} into Φ^i or into a subcomplex of Φ^i. We call τ_i a *projection* of Φ^{i+1} onto Φ^i, and more generally $\tau_i\tau_{i+1} \cdots \tau_{i+j-1}\Phi^{i+j}$ a projection of Φ^{i+j} onto Φ^i. The latter is also a simplicial transformation of Φ^{i+j} into Φ^i or into a subcomplex of Φ^i.

3. I say that in fact $\tau_i\Phi^{i+1} = \Phi^i$, that is every simplex of Φ^i is the transform of a simplex of Φ^{i+1}, or, in other words, Φ^i is completely covered by $\tau_i\Phi^{i+1}$. For let us suppose that $\tau_i\Phi^{i+1} = \Psi$, a proper subcomplex of Φ^i. There exists then a simplex $\sigma_p = U^{i\alpha_0} \cdots U^{i\alpha_p} \subset \Phi^i - \Psi$. Denote generically by V^α the sum of all the sets $U^{i+1,\beta}$ which make up $\tau_i^{-1}U^{i\alpha}$; clearly $V^\alpha \subset U^{i\alpha}$. Since every U^{i+1} corresponds to one (and only one) V, $\Sigma = \{V^\alpha\}$ is an ϵ_i-f. c. o. s. and it has a subcomplex Φ' of Φ as its skeleton. I say that σ_p is not a cell of Φ'. For otherwise we would have $V^{\alpha_0} \cdots V^{\alpha_p} \neq 0$, and hence there would exist a $U^{i+1,\beta_0} \cdots U^{i+1,\beta_p} \neq 0$, where U^{i+1,β_h} is a constituent of V^{α_h}. Since $\tau_i U^{i+1,\beta_h} = U^{i\alpha_h}$, we would then have in $\sigma'_p = U^{i+1,\beta_0} \cdots U^{i+1,\beta_p}$ a simplex of Φ^{i+1} such that $\tau_i\sigma'_p = \sigma_p$ and hence $\sigma_p \subset \Psi = \tau_i \cdot \Phi^{i+1}$, contrary to assumption.

Under the circumstances then $\sigma_p \not\subset \Phi'$. It follows that Φ' is a proper subcomplex of Φ^i and also the skeleton of an ϵ_i-f. c. o. s. But this is ruled out since Σ^i is irreducible. Hence σ_p cannot exist, and $\Psi = \tau_i\Phi^{i+1} = \Phi^i$.

Definitions. A sequence $\{B^i\}$ of elements, (sets, complexes, etc.) such that $B^i \subset \Phi^i$ and $\tau_i B^{i+1} = B^i$ is called a *projection-sequence* (of sets, of complexes, etc.).

Given any (non-singular) chain C_p we shall designate by $| C_p |$ the complex made up of the cells of the chain. A sequence $\{C_p{}^i\}$ will be called a projection-sequence of chains or cycles whenever

$$\tau_i C_p{}^{i+1} = C_p{}^i.$$

4. Let $U^{i\delta} \supset x$. There exists a set $U^{i-1,\gamma} \supseteq U^{i\delta}$ such that $\tau_{i-1}U^{i\delta} = U^{i-1,\gamma}$; a set $U^{i-2,\beta}$ similarly related to $U^{i-1,\gamma}$, etc., clear up to a certain set U^{1a}. Let k_i be for each i the class of all sets U^{1a} thus obtained. The classes k_i are all finite, $\neq 0$ and $k_i \supset k_{i+1}$. Therefore from a certain i on $k_i = k_{i+1} = \cdots$. Consequently there exists an infinite sequence $\{U^{ia_i}\}$ such that $U^{ia_i} \supseteq U^{i+1,a_{i+1}}$, $\tau_i U^{i+1,a_{i+1}} = U^{ia_i}$, $\Pi U^{ia_i} = x$. Let $U^{ia_0}, \cdots, U^{ia_p}$ be all the sets of Σ^i occurring in any such sequence corresponding to the same point x and let $V^i = U^{ia_0} \cdots U^{ia_p} \neq 0$, so that $\sigma_p{}^i = U^{ia_0} \cdots U^{ia_p}$ is a simplex of Φ^i. Since every U^{ia} here occurring is the τ_i transform of a similar $U^{i+1,\beta}$ we have $\tau_i\sigma^{i+1} = \sigma^i$, hence $\{\sigma^i\}$ is a projection-sequence of simplexes. Moreover clearly $V^i \supseteq V^{i+1}$, $\Pi V^i = x$.

Conversely if $\{\sigma^{*i}\}$ is a projection-sequence of simplexes, and if V^{*i} is the intersection of the sets U^i associated with σ^{*i}, then $V^{*i} \supseteq V^{*i+1} \supset x$, hence $\Pi V^{*i} = x$. Clearly also the sets U^i associated with σ^{*i} are among those associated with σ^i, hence σ^{*i} is σ^i or a face of σ^i. We call $\{\sigma^{*i}\}$ and $\{\sigma^i\}$ respectively *projection-sequence* and *maximal projection-sequence* for the point x.

5. Owing to the choice of $\{\Sigma^i\}$ we may use $\{\Phi^i\}$ to map the space $\mathfrak{R}$ topologically on an Euclidean S_r, $r \geqq 2n + 1$.† Choosing $r \geqq 2n + 2$ we may even carry out the mapping so as to be able to construct the joining cell of any simplex of Φ^{i+1} with its transform under τ_i (deformation cell corresponding to τ_i), and from there, as the sum of all these cells, the $(n+1)$-complex K or *fundamental complex* of $\mathfrak{R}$, (*Topology*, p. 327) which will be an infinite complex on S_r. The part of K obtained on removing Φ^i, $\Phi^{i-1}, \cdots$ and the cells joining them will be denoted by N^i and the finite complex $K - N^i$ by K^i.

In practice we shall find it more convenient to have a representation of $\mathfrak{R}$ and K on the Hilbert parallelatope

$$\mathfrak{H}: \quad 0 \leqq x_i \leqq 1/i, \qquad\qquad (i = 1, 2, \cdots + \infty).$$

This image is to be constructed as follows. As proved by Urysohn $\mathfrak{R}$ has a topological image $\mathfrak{R}'$ on $\mathfrak{H}$. Consider now the following homeomorphism of $\mathfrak{H}$:

$$T: \quad x'_i = \tfrac{1}{2}(x_i + 1/i)$$

which transforms it into the subset

$$\mathfrak{H}': \quad 1/2i \leqq x_i \leqq 1/i.$$

Then $T\mathfrak{R}' = \mathfrak{R}''$ is a topological image of $\mathfrak{R}$ which possesses no point for which any x_i is zero. We identify henceforth $\mathfrak{R}$ with $\mathfrak{R}''$.

† See our paper in the *Annals of Mathematics*, Vol. 32 (1931), p. 528.

Let us denote by S^i the subset of $\mathfrak{H}$ consisting of all points for which $x_k = 0$ when $k > (2n + n_i)$, where n_i increases so rapidly that we may carry out the construction of K, given in *Topology*, p. 325, in such manner that $\Phi^i \subset S^i - S^{i-1}$. As a consequence $K \cdot \mathfrak{R} = 0$. Now, with closures referring to $\mathfrak{H}$, the only limit-points of $\bar{K}$ not on K are on $\mathfrak{R}$, hence $\mathfrak{R} = \bar{K} - K$. It is in order to fulfill this condition that the complex K has been constructed in the above special manner.

6. It is convenient to join each point of Φ^{i+1} to its transform by τ_i by a segment in $\mathfrak{H}$. The sum of these segments coincides with K. An infinite arc consisting of a sequence of projecting segments for $\tau_1, \tau_2, \cdots$ plus their coterminal end-points, will be called a *projecting line*. The projecting lines all start at Φ^1, which we designate henceforth by Φ, and continue indefinitely throughout K.

If $\{B^i\}$ is a projection sequence of sets or complexes, the set $\mathfrak{B}$ obtained by adding to the sequence the projecting segments of the points of the B's is called a *projection-set*. If the B's are complexes, the projecting segments of a definite p-cell of B^{i+1} make up a $(p+1)$-cell; these are the joining cells of B^{i+1} and $\tau_i B^{i+1}$ (No. 5). The sum of the closures of all these cells is a *projection-complex* $\mathfrak{K}$. If B^i is a subcomplex of Φ^i for every i, $\mathfrak{K}$ is a subcomplex of K.

We are primarily interested in the relation between various subcomplexes of K and certain associated sets of $\mathfrak{R}$. Properly speaking instead of a subcomplex of K we might well take any subset of K, but actually the subcomplexes will suffice for our purpose.

With any subcomplex L of K we may associate the closed subset $F = \bar{L} \cdot \mathfrak{R}$, and we observe immediately that this set F depends solely upon the "infinite" part of L, i. e. it is unchanged when a finite complex is added to or removed from L. In the sense of *Topology*, Ch. VII, $\mathfrak{R}$ is associated with the total ideal element of K, and F with a certain closed ideal element of the complex.

Suppose that we construct a new fundamental K' for $\mathfrak{R}$, that we suppose as before on $\mathfrak{H}$, and such that $K' \cdot \mathfrak{R} = 0$. Applying to K the deformation theorem of *Topology*, p. 328 † (proved for chains but applicable to complexes), we can reduce L to a subcomplex L' of K' by a deformation that $\to 0$ for any particular cell of K as that cell $\to \mathfrak{R}$. Therefore $F = \bar{L}' \cdot \mathfrak{R} = \bar{L} \cdot \mathfrak{R}$, i. e. the set F is in a large measure independent of the complex K.

† In the proof *loc. cit.*, A_j should be mapped on $\tau_{i-1}\lambda_i A_j$. Owing to the condition $\epsilon_i < \frac{1}{2}\eta_{i-1}$ which we have imposed, $C_p{}^{i-1}$ will still be mapped as before on a subchain of Φ^i.

7. We shall now reverse the situation: starting with any particular closed set F we shall associate with it a certain projection-complex L, such that $F = \bar{L} \cdot \mathfrak{R}$ and $\dim F = p = \dim L - 1$, which is the maximum value possible for p.

According to Menger (*Dimensionstheorie*, p. 158) there is a f. c. o. s. of order $\leqq p$ of F (not of $\mathfrak{R}$), $\Sigma'^i = \{V^{ia}\}$, such that there is one and only one V^{ia} on any U^{ia} that meets F. When i exceeds a certain value the skeleton Φ'^i of Σ'^i is a p-complex. Associate with each V^{ia} the vertex U^{ia} of the set of same name. Now when a certain aggregate of sets V^i intersect, the same holds as regards the corresponding sets U^i. Hence Φ'^i will thus become a subcomplex of Φ^i. Now take all the subcomplexes Ψ'^1 of Φ^1 which are the projections of a Φ'^i. Since their number is infinite and the number of subcomplexes of Φ^1 is finite, at least one, Ψ^1 is the projection of an infinity of complexes Φ'^i. Consider the subcomplexes Ψ'^2 of Ψ^2 such that $\tau_1\Psi'^2 = \Psi^1$. There is an infinity of complexes Φ'^i, $i \geqq 2$, projected onto Ψ^1 and their projections on Φ^2 are each a Ψ'^2. Therefore at least one of the latter, Ψ^2, is the projection of an infinity of complexes Φ'^i, etc. By this obvious process we obtain an infinite projection-sequence $\{\Psi^i\}$, where Ψ^i is a subcomplex of Φ^i which is the projection of a Φ'^j, and $\dim \Psi^i \leqq \dim \Phi'^j \leqq p$. Since Φ'^j is the skeleton of an ϵ_j-f. c. o. s. of F, the latter may be $6\epsilon_j$-deformed into F.† Moreover, referring to the representation in $\mathfrak{H}$, Φ'^j can be ξ_i-deformed into Ψ^i, ($\xi_i \rightarrow 0$ with $1/i$). Hence F can be ζ_i-deformed into Ψ^i, ($\zeta_i \rightarrow 0$ with $1/i$). Therefore Ψ^i is the skeleton of a θ_i-f. c. o. s. of $F(\theta_i \rightarrow 0$ with $1/i)$ (Alexandroff, *loc. cit.*, p. 18). As a consequence if we put in the joining cells of the Ψ's, we obtain a fundamental complex L for F. We have $\dim L = p + 1$, for it is $\geqq p + 1$ since $\dim F = p$, and $\leqq p + 1$, since $\dim \Psi^i \leqq p$.

Since we have but little information regarding the meshes or the characteristic constants of the coverings of F whose skeleta are the Ψ's, it is not easy to show that the deformation theorem applies to L. Therefore for the homology theory another similar $(p + 1)$-complex L^* is more suitable. It is constructed as follows: take the skeleton of the aggregate $\{U^{ia} \cdot F\}$ (i fixed) and remove from it all cells of dimension $> p$. What is left is a subcomplex Ω^i of Φ^i, and we have immediately, owing to the mode of constructing the Φ's, $\tau_i\Omega^{i+1} \subset \Omega^i$. The complex L^* consists of all the Ω's plus their joining cells. It is clearly a $(p + 1)$-subcomplex of K, which we shall call the *generalized fundamental complex* of the set F. The proof of the deformation theorem is directly applicable to L^* for all cycles or complexes of dimension $\leqq p$. Since

† P. Alexandroff, *Annals of Mathematics*, Vol. 30 (1928-29), p. 13.

$\dim F = p$, F possesses a fundamental $(p+1)$-complex L' to which the deformation theorem is applicable. For example L' may be built up out of a subset of the Ψ's. It follows (see No. 9), that the q-cycles, $q > p$, of F are all $\equiv 0$ and hence they need not concern us further.

8. *The chains and cycles of K.* The only chains of K with which we shall be concerned are its subchains, no others being considered. Whatever C_p we have: $C_p = C'_p + C_p''$, where C'_p is the part of C_p on K^i and C_p'' the rest. It is convenient to write: $C'_p = K^i \cdot C_p$, $C_p'' = N^i \cdot C_p$. The part of $F(C'_p)$ which is on Φ^i will be designated by $\Phi^i \cdot C_p$ and called the *trace* of C_p on Φ^i.

Let us suppose that we have on K an aggregate of chains $\{C_q^i\}$, $q = 0, 1, \cdots, p$; $i = 1, 2, \cdots$, such that: (a) $C_p = \Sigma C_p^i$ is a true chain of K, i. e. includes no cell of K taken with an infinite coefficient; (b) for every C_q^i, we have

$$C_q^i \to \Sigma \eta^q_{ij} C^j_{q-1}. \tag{8.1}$$

The aggregate $\{C_q^i\}$ is called an *elementary decomposition* of C_p. An example is of course the decomposition of C_p into its cells. For later purposes a more general decomposition is introduced here.

Two decompositions $\{C_q^i\}$, $\{C'_q{}^i\}$ of two chains C_p, C'_p are said to have the same structure if they correspond to one another chain for chain (for every C_q^i one and only one chain $C'_q{}^i$ and conversely) and if the corresponding incidence numbers η^q_{ij} are the same. That is to say if the sets are labelled in such manner that C_q^i and $C'_q{}^i$ are the associated chains in the correspondence then they have the same incidence matrices $\| \eta^q_{ij} \|$.

Suppose now that we have two decompositions $\{C_q^i\}$, $\{C'_q{}^i\}$ of C_p, C'_p whose structure is the same, and let there exist for every C_q^i a $(q+1)$-chain $\mathfrak{D} C_q^i$, called a *deformation-chain*, such that

$$\mathfrak{D}\, C_q^i \to C'_q{}^i - C_q^i - \Sigma \eta^q_{ij} \mathfrak{D} C^j_{q-1}. \tag{8.2}$$

If we agree to write

$$\mathfrak{D} \Sigma\, a_i C_q^i = \Sigma\, a_i \mathfrak{D}\, C_q^i \tag{8.3}$$

then (8.2) assumes the form

$$\mathfrak{D}\, C_q^i \to C'_q{}^i - C_q^i - \mathfrak{D} F(C_q^i). \tag{8.4}$$

Under the circumstances the passage from C_p to C_p' is called a *deformation* of C_p into C'_p, and $\mathfrak{D}\,(C_p)$ is called the *deformation-chain* of C_p.

A deformation of a subcomplex L of K into another L' could be defined substantially along similar lines. We would merely replace the chains C_q^i by

the cells $E_q{}^i$ of L, and in (8.2), (8.3), (8.4), the C's would be cells and the $\mathfrak{D}C$'s would continue to be chains but otherwise the rest would be as before. Then $\Sigma \mid \mathfrak{D}E \mid + L + L'$ would be called the *deformation-complex* $\mathfrak{D}L$ of L.

All this is entirely in line with the treatment of deformations in *Topology*, p. 78, except that there we had only cells and obtained (8.2) from direct geometric considerations, essentially by considering the deformation as a "singular" translation, whereas (8.2) serves directly to define the deformation. This departure is justified on the ground that (8.2) is the central property of a deformation as regards the applications to any homology theory.

For purposes of reference, if we agree to neglect everywhere chains on L, or else if we only consider integral chains mod m or both we have associated deformations and deformation-chains mod L, mod m, mod (L, m), as the case may be.

If in a given deformation $\mathfrak{D}$ every deformation chain is of diameter $< \epsilon$ we have a so-called *ϵ-deformation.*

By analogy with ordinary deformations we shall say that $\mathfrak{D}$ leaves a chain $C_q{}^i$ *invariant* or *does not displace the chain,* whenever the chain $\mathfrak{D}C_q{}^i = 0$.

9. *The chains and cycles of the space $\mathfrak{R}$.* Taking substantially the point of view of *Topology*, Ch. VII, § 4, we consider a $(p+1)$-chain C_{p+1} of K as defining a p-chain c_p of $\mathfrak{R}$, a cycle mod Φ, Γ_{p+1} of K as defining an absolute p-cycle γ_p of $\mathfrak{R}$. In particular if

$$K \supset C_{p+2} \rightarrow \Gamma_{p+1} \quad \text{mod } \Phi,$$

and if C_{p+2} determines c_{p+1} of $\mathfrak{R}$ we write

$$c_{p+1} \rightarrow \gamma_p, \qquad \gamma_p = F(c_{p+1}),$$

and say "γ_p bounds c_{p+1}". A special case is where Γ_{p+1} is finite, for it is then ≈ 0 mod Φ, since it can be deformed along the projecting lines onto Φ. We say that γ_p is homologous to zero: $\gamma_p \approx 0$, whenever it is a finite or infinite sum of bounding cycles. The extension to cycles mod A, A closed, is in the usual manner: Γ_{p+1} is then a cycle mod L, where L is any subcomplex of K such that $\bar{L} \cdot \mathfrak{R} = A$.

The p-th homology group $\mathfrak{H}_p$ (absolute or mod A) is the quotient group $\mathfrak{G}_p \div \mathfrak{G}'_p$ of the Abelian group $\mathfrak{G}_p$ of the *p-cycles* (written additively) by the group $\mathfrak{G}'_p$ of the cycles ≈ 0. The bases and homology characters are defined as usual.

For $p = n$ the bounding relations between the cycles are reduced to the identical linear relations between them. In terms of the n-cycles it is possible to define the generalized absolute orientable n-circuit (*Topology*, p. 76): it is

a compact metric n-space $\mathfrak{R}$ such that $R_n(\mathfrak{R}) = 1$, and $R_n(A) = 0$ for any *proper* closed subset A of $\mathfrak{R}$. As a consequence the circuit has a base for the n-cycles consisting of a single γ_n, i. e. every n-cycle of $\mathfrak{R}$ is of the form $t\gamma_n$. In place of γ_n we might as well take $-\gamma_n$ and either one of the pairs $(\mathfrak{R}, \gamma_n)$, $(\mathfrak{R}, -\gamma_n)$ is called an *oriented* circuit, the passage from one to the other being described as a reversal of orientation. The non-orientable circuit is obtained by taking the cycles mod 2, and similarly for the circuits mod m. Analogous notions hold for the circuit mod A, A closed, the circuit conditions being $R_n(\mathfrak{R}, A) = 1$, $R_n(B, A) = 0$, where B is now any proper closed subset of $\mathfrak{R}$ which $\supset A$.

10. We have taken the chains and cycles of $\mathfrak{R}$ as represented by actual chains or cycles mod Φ of K. Their characteristic part corresponds however to the infinite portion of the representative C_{p+1} or Γ_{p+1}. As a matter of fact the difference is not great: we may always suppress, say $\Phi^1, \cdots, \Phi^k$, with all the cells joining them, and consider Φ^k as the new Φ, thus converting any Γ with a finite boundary into a cycle mod Φ. Another way of looking at the matter is as follows: under our conventions for chains the suppression of any finite part of C_{p+1} is not to affect c_p. As for a γ_p it is then to be represented by a C_{p+1} with finite boundary C_p. But if we slide the points of C_p along the projecting lines down onto Φ, and add the deformation-chain, which is finite, to C_{p+1}, we have a cycle mod Φ, Γ_{p+1}, which also represents γ_p.

The set $\mathfrak{R} \cdot |\overline{C_{p+1}}|$, where as before the closure refers to $\mathfrak{H}$, is a closed subset of $\mathfrak{R}$ associated with c_p, that we shall denote by $|c_p|$. This set depends solely on c_p, and not on the particular fundamental complex K chosen (No. 6).†

By the points of c_p we shall always mean the points of $|c_p|$. In particular a set A is said to intersect c_p whenever it intersects $|c_p|$, to be $\subset c_p$ or to $\supset c_p$ whenever $A \subset |c_p|$ or $\supset |c_p|$ as the case may be.

Let A be a closed set. By a p-cycle mod A we shall mean a c_p such that $F(c_p) \subset A$. The cycle is said to *bound* mod A whenever there exists a c_{p+1} such that $F(c_{p+1}) - c_p \subset A$. Finally it is ≈ 0 mod A whenever the cycle is a finite or infinite sum of cycles which bound mod A.

We may also consider the absolute cycles of $\mathfrak{R} - A$. Such a cycle is ≈ 0 on $\mathfrak{R} - A$ whenever it is ≈ 0 on some closed subset of $\mathfrak{R} - A$.

† The p-chains such that $\dim |c_p| \leqq p$ form a topological subclass of the class of all p-chains. These special chains played an important part in the initial version of the present paper. We found it simpler since then, to eliminate them entirely, and to replace them everywhere merely by the projection-chains which are introduced in No. 13. As the properties needed in Part II were only those of projection-chains, the only important modifications required were in Nos. 11, 18, 19 (June, 1933).

11. A deformation of a C_{p+1} into C'_{p+1} on K may serve to define two kinds of deformations $\mathfrak{D}$ of the associated chains c_p, c'_p on $\mathfrak{R}$. The deformation $\mathfrak{D}$ is of the *first kind* whenever the chains of the associated elementary decompositions $\{C_q{}^i\}$, $\{C'_q{}^i\}$, are all finite; it is of the *second kind* when some or all are infinite.

Consider for the present a $\mathfrak{D}$ of the first kind. If the deformation-chain of $C_q{}^i \to 0$ with $1/i$, we consider the two chains c_p, c'_p as identical. If U is any open set $\supset c_p$, and if L is any subcomplex of K such that $\overline{L} \cdot \mathfrak{R} = \bar{U}$, then for i above a certain value $\mathfrak{D}\, C_q{}^i \subset L$, and hence C_{p+1} has at most a finite subchain on $K - L$.

As an application if C_{p+1} is deformed over $\mathfrak{H}$, according to the deformation theorem of *Topology*, p. 328, into a new chain C'_{p+1} of K, then the chain c'_p defined by C_{p+1} is identical with c_p. For the deformation over $\mathfrak{H}$ gives rise to a certain deformation-chain $\mathfrak{D}\, C_{p+1}$ with a suitable elementary decomposition. If we now reduce $\mathfrak{D}\, C_{p+1}$ to K by the deformation theorem, choosing, as we may, the chains of the decompositions which it demands (the analogues of the chains $C_p{}^i$ of the proof *loc. cit.*) exact sums of chains of the decomposition of C_{p+1}, the sole effect of the deformation on C_{p+1}, C'_{p+1} may be to subdivide them, and this has no influence on c_p, c'_p. As a consequence we have on K a deformation-chain for a deformation of C_{p+1} into C'_{p+1} which is of the first kind. Hence $c_p \equiv c'_p$.

Suppose in particular that we have a closed set A with L^*_A as its generalized fundamental complex (No. 7) and let γ_p be a cycle mod A. If Γ_{p+1} is the representative chain of γ_p, $F(\Gamma_{p+1})$ represents the absolute cycle $F(\gamma_p)$ of A. This absolute cycle has a representative image Γ'_p which is a cycle of L^*_A mod Φ (No. 7) and by the above

$$K \supset D_{p+1} \to \Gamma'_p - F(\Gamma_{p+1})\,;$$
$$\Gamma'_{p+1} = \Gamma_{p+1} + D_{p+1} \to \Gamma'_p; \quad \overline{|D_{p+1}|} \cdot \mathfrak{R} \subset A.$$

Hence if Γ'_{p+1} represents γ'_p of $\mathfrak{R}$ we have $\gamma'_p - \gamma_p \subset A$ so that γ'_p represents the same cycle mod A as γ_p. Therefore we may represent a cycle mod A by a chain C_{p+1} whose boundary is on the generalized fundamental complex L^*_A of the set A. This result will be useful later.

The only deformations occurring in the sequel are of the second kind, and the elementary decompositions and deformation-chains on K will always be in finite number. This will be understood throughout. They determine elementary decompositions $\{c_q{}^i\}$, $\{c'_q{}^i\}$, and deformation-chains $\mathfrak{D}\, c_q{}^i$ for the deformation of c_p into c'_p, and the rest is as in No. 8. In particular

$$\mathfrak{D}\, c_p \to c'_p - c_p - \mathfrak{D} F(c_p)\,; \tag{11.1}$$

(11.2) $$\mathfrak{D}\gamma_p \approx \gamma'_p - \gamma_p \approx 0 \text{ on } \mathfrak{R}.$$

12. With notations as in No. 11, let γ_p be a cycle mod A whose representative Γ_{p+1} has its boundary on $L^*_A + \Phi$. The N S C in order that $\gamma_p \approx 0 \bmod A$, is that for every i

(12.1) $$\Gamma_{p+1} \approx 0 \bmod (N^i + \Phi + L^*).$$

Whether the cycle is ≈ 0 or not when (12.1) holds for any particular i it holds also for the lower values of i. Therefore there is an h, called the *index* of γ_p, such that (12.1) holds for $i \leqq h - 1$ but not for $i \geqq h$. It implies that there exists an infinite cycle $\Gamma'_{p+1} \subset N^{h-1}$ such that

(12.2) $$\Gamma_{p+1} \approx \Gamma'_{p+1} \bmod \Phi,$$

while no such cycle exists for any N^i, $i \geqq h$. In terms of the traces we have at once

(12.3) $$\Phi^i \cdot \Gamma_{p+1} \approx \Phi^i \cdot \Gamma'_{p+1} \text{ on } \Phi^i, \quad (i \geqq h),$$

(12.4) $$\Phi^i \cdot \Gamma_{p+1} \approx 0 \text{ on } \Phi^i, \quad (i < h).$$

Conversely suppose that (12.4) holds for $i < h$ but not for any higher i. We have then

(12.5) $$\Phi^i \supset D_{p+1} \rightarrow \Phi^i \cdot \Gamma_{p+1},$$

(12.6) $$D'_{p+1} = N^i \cdot \Gamma_{p+1} - D_{p+1} \rightarrow 0.$$

Since the cycle D'_{p+1} is finite it is $\approx 0 \bmod \Phi$ on K, for it can be projected onto Φ. It follows that (12.2) holds with $\Gamma'_{p+1} = \Gamma_{p+1} - D_{p+1} \subset N^{i-1}$ and hence the index $\geqq h$. On the other hand the index $\leqq h$, since otherwise (12.4) would hold for some $i \geqq h$. Therefore *the index h of Γ_{p+1} is the highest value of $i + 1$ for which* (12.4) *holds.*

13. We may consider τ_i as a deformation of Φ^{i+1} into Φ^i over K. The cell joining E_p of Φ^{i+1} with $\tau_i E_p$, suitably oriented, is the deformation-chain of E_p (*Topology*, p. 78), and the deformation-chain of any subchain C_p^{i+1} of Φ^{i+1} is then obtained as *loc. cit.* by the condition that it is a linear chain-function. If we designate this function by $\mathfrak{D}$ we have

(13.1) $$\mathfrak{D} C_p^{i+1} \rightarrow C_p^i - C_p^{i+1} - \mathfrak{D} F(C_p^{i+1}),$$

(13.2) $$C_p^i = \tau_i C_p^{i+1}.$$

If k^{i+1} is any subcomplex of Φ^{i+1} the sum of the closed deformation-cells of its cells (deformation-chains of the cells) under τ_i is a complex $\mathfrak{D} k^{i+1}$, the

deformation-complex of k^{i+1}. If we have an infinite sequence of complexes $\{k^{i+1}\}$, where k^{i+1} is a subcomplex of Φ^{i+1} and $k^i = \tau_i k^{i+1}$ for every i, the sum $k = \Sigma \mathfrak{D} k^{i+1}$ is a *projection-complex.*

Let now $\{C_p^{i+1}\}$ be an infinite sequence of chains where C_p^{i+1} is a subchain of Φ^{i+1} and $C_p^i = \tau_i C_p^{i+1}$ for every i. We have then an associated chain

$$C_{p+1} = \Sigma \mathfrak{D} C_p^i \tag{13.3}$$

defining a chain c_p of $\mathfrak{R}$, called a *projection-chain.* If the chains C_p^i for i above a certain value h are cycles Γ_p^i, C_{p+1} defines a γ_p called a *projection-cycle* of $\mathfrak{R}$. The chains C_p^j, $j \leqq h$, can be replaced by the projections of Γ_p^{h+1} without modifying γ_p, so that when we have a γ_p we may assume that all the chains C_p^i are cycles.

Let $\{C_p^i\}$ define as above a projection-chain c_p, with C_{p+1} as the associated chain of K. Then $| C_p^i |$ is not necessarily the projection of the complexes $| C_p^{i+j} |$, but their difference is made up of cells of less than p dimensions. It follows that there exists a projection-complex k such, that for each i, $k^i - | C_p^i |$ consists of cells of dimension $< p$, while k^i is the projection of some C_p^{i+j} on Φ^i. The difference $k - C_{p+1}$ will consist of cells of dimension $< p + 1$.

Let A, L^*_A be as before and let C_{p+1} be any chain of K with C_p^i as its traces. We may introduce as above the finite chains $\mathfrak{D} C_p^i$ and also the infinite chain

$$C'_{p+1} = C_{p+1} - \Sigma \mathfrak{D} C_p^i.$$

Let C_{p+1} define c_p of $\mathfrak{R}$. Whenever $C'_{p+1} \subset L^*_A$ we shall call C_{p+1} a *projection-chain mod* L^*_A, and c_p a *projection-chain mod* A, a *projection-cycle mod* A when $F(c_p) \subset A$. When $A = 0$ we have $L^*_A = 0$, $C'_{p+1} = 0$ and c_p becomes an ordinary projection-chain.

14. *Certain properties of chain-moduli.* By a *modulus* of p-chains of a complex K we understand a system of rational chains of K forming an abelian group with respect to addition. If $\mathfrak{M}, \mathfrak{N}$ are two such moduli, and if $\mathfrak{N} \subset \mathfrak{M}$ then, as usual, $C_p \equiv 0 \bmod \mathfrak{N}$ or $C_p \equiv C'_p \bmod \mathfrak{N}$, mean that $C_p \subset \mathfrak{N}$ or $C_p - C'_p \subset \mathfrak{N}$. If $\mathfrak{M}'$ is a submodulus of $\mathfrak{M}$, by $\mathfrak{M} \equiv \mathfrak{M}' \bmod \mathfrak{N}$, we shall mean that every element of $\mathfrak{M}$ is congruent to an element of $\mathfrak{M}'$ mod $\mathfrak{N}$ and conversely. If we have a modulus $\mathfrak{M}$ on Φ^j the projections of its chains on Φ^i, $i < j$, constitute a modulus $\mathfrak{M}'$ called the *projection* of $\mathfrak{M}$ on Φ^i. Regarding these moduli of the Φ's and their projections we shall prove the following important

THEOREM I. *Let there be given for every h two moduli of p-chains of Φ^h,*

$\mathfrak{M}^h$ and $\mathfrak{N}^h \subset \mathfrak{M}^h$, such that the projection of any $\mathfrak{M}^j$, $j \geqq h$, on Φ^h, is congruent to $\mathfrak{M}^h$ mod $\mathfrak{N}^h$. Then corresponding to every C_p of $\mathfrak{M}^h$, there is a projection-sequence $\{C_p{}^i\}$, $C_p{}^i \subset \mathfrak{M}^i$, such that $C_p \equiv C_p{}^h$ mod $\mathfrak{N}^h$.

If $C_p \subset \mathfrak{N}^h$ we may take a vacuous sequence as the corresponding $\{C_p{}^i\}$. Therefore we may assume that $C_p \not\subset \mathfrak{N}^h$. Under the circumstances C_p is a proper p-chain and so is any chain $C'_p \equiv C_p$ mod $\mathfrak{N}^h$.

Consider then all the chains of $\mathfrak{M}^h$, $D_p \equiv C_p$ mod $\mathfrak{N}^h$, which are projections of chains of some $\mathfrak{M}^j$, $j \geqq h$. The number of projections being infinite and the number of subcomplexes $| D_p |$ of Φ^h finite, at least one of these subcomplexes must carry an infinity of chains D_p. Let $\mathcal{K}$ be such a $| D_p |$ with the least number possible, s, of p-cells and let $E_p{}^1, \cdots, E_p{}^s$ be its p-cells, so that

$$D_p = \Sigma\, t_a E_p{}^a.$$

The chain D_p is the projection of an element say of $\mathfrak{M}^j$. Suppose that there is another similar chain

$$D'_p = \Sigma\, t'_a E_p{}^a$$

which is the projection of an element of $\mathfrak{M}^k$, $k \geqq j$. Then D'_p is likewise the projection of an element of $\mathfrak{M}^j$ and

$$D_p'' = \Sigma(t_a - t'_a) E_p{}^a = \Sigma\, t_a'' \cdot E_p{}^a$$

is the projection of an element of $\mathfrak{M}^j$ which is in $\mathfrak{N}^h$. Therefore, if, no matter how high we take j, there are in $\mathfrak{M}^j$ two elements whose projections D_p, D'_p are different, and both $\equiv C_p$ mod $\mathfrak{N}^h$, there exists always in $\mathfrak{M}^j$ an element whose projection D_p'' is a chain of $\mathcal{K}$ and in $\mathfrak{N}^h$.

Conceivably some, but not all the t'''s vanish for j high enough. There will be one, however, say $t_1'' \neq 0$ for an infinity of j's, hence for every $j > h$, and $D_p - t_1 D_p''/t_1''$ will be a subchain of $\mathcal{K} - E'$ which is $\equiv C_p$ mod $\mathfrak{N}^h$. We have thus a complex whose number of p-cells $< s$, and which carries an infinity of chains such as D_p. As this contradicts the assumption regarding s, it follows that for j above a certain value $D_p'' \equiv 0$, $D_p \equiv D'_p$. Therefore there is a unique chain of $\mathcal{K}$ which is the projection of chains of $\mathfrak{M}^j$, j above a certain value, and $\equiv C_p$ mod $\mathfrak{N}^h$.

Let us now write $D_p{}^h$ for D_p and consider the chains of $\mathfrak{M}^{h+1}$ whose projection on Φ^h is $D_p{}^h$. Their number being clearly infinite, we may again choose one, $D_p{}^{h+1}$, consider the least number possible of its p-cells and show that if it is not unique $D_p{}^h$ can be replaced by a similar chain with a smaller s, etc. We thus obtain a sequence $\{D_p{}^i\}$ $i = h, h+1, \cdots$. The projection sequence

$\{C_p{}^i\}$ such that $C_p{}^i = D_p{}^i$ for $i \geqq h$; $C_p{}^{h-i} =$ the projection of $D_p{}^h$ on Φ^{h-i}, has all the properties required by Theorem I.

15. *Remarks.* I. In the proof the fact that the chains are taken with rational coefficients enters in an essential manner when we multiply chains by the number s_α/r_α. *Clearly any ring of coefficients forming a field (i. e. with unique division) would be admissible,* for instance the ring of integers mod p, p a prime. But we could not have integers mod m, m not a prime.

II. Let us call a projection-sequence $\{C_p{}^i\}$, $C_p{}^i \subset \mathfrak{M}^i$, *irreducible* whenever for any other similar $\{C'_p{}^i\}$, such that $C'_p{}^i$ is a subchain of $|\, C_p{}^i \,|$, necessarily $C'_p{}^i = tC_p{}^i$. In that case of course t is independent of i. If we examine our construction we see that the sequence $\{C_p{}^i\}$ of our theorem has been chosen irreducible. For the irreducibility condition is imposed when $i \geqq h$, and follows, by projection, when $i < h$.

16. If we consider again the elements of $\mathfrak{M}^h$ where h is now fixed, I say that *we can construct for $\mathfrak{M}^h$ a finite base mod $\mathfrak{N}^h$, $C_p{}^{h\alpha}$, $\alpha = 1, 2, \cdots, r$, whose elements are members of irreducible projection-sequences $\{C_p{}^{i\alpha}\}$.*

Let $E_p{}^\beta$ denote this time all the p-cells of Φ^h. By the procedure of *Topology*, p. 302 (method of the "first-cell,") and with a suitable numbering of the cells, Theorem II authorizes us to assume that, except for irreducibility, we already have the required base such that in addition

$$C_p{}^{h\alpha} = E_p{}^\alpha + \Sigma\, t_{\alpha\gamma} E_p{}^{r+\gamma}. \tag{16.1}$$

Consider now the subcomplex Ψ^i of Φ^i consisting of all the cells of Φ^i projected onto $|\, C_p{}^{h\alpha} \,|$ and apply the theorem to $\{\Psi^i\}$ taking as modulus $\mathfrak{M}^{*i}$ the aggregate of the elements of $\mathfrak{M}^i$ that are subchains of Ψ^i. Since the elements of $\mathfrak{M}^{*h}$ all are, mod $\mathfrak{N}^h$, linear combinations of the chains $C_p{}^{h\alpha}$, and contain only $E_p{}^\alpha$ among the first r p-cells, they are all, mod $\mathfrak{N}^h$, multiples of $C_p{}^{h\alpha}$, and those in $\mathfrak{M}^h - \mathfrak{N}^h$ must contain $E_p{}^\alpha$. Now by Th. I taken together with No. 15 Remark II, we can find precisely an irreducible $\{C_p{}^{i\alpha}\}$ such that $C_p{}^{h\alpha}$ is of the form (16.1), and in particular congruent mod $\mathfrak{N}^h$ to the chain $C_p{}^{h\alpha}$ in (16.1). Therefore $\{C_p{}^{i\alpha}\}$, $\alpha = 1, 2, \cdots, r$, behaves as required.

17. Consider the $(p+1)$-subchains C_{p+1} of K, such that $C_{p+1} \cdot \Phi^h \subset \mathfrak{M}^h$ for every h. Thus if $\{C_p{}^i\}$ are the projection-sequences that we have just considered the closures of the joining cells of the chains $C_p{}^i$ form a chain such as C_{p+1}. The finite or infinite linear combinations of these chains which are chains form a modulus $\mathfrak{M}$ and the similar chains corresponding to the moduli $\mathfrak{N}^h$ form a submodulus $\mathfrak{N}$ of $\mathfrak{M}$. We shall say that any particular projection-

chain C_{p+1} of $\mathfrak{M}$ is *irreducible* if it contains no similar subchain (member of $\mathfrak{M}$) which is not a multiple of C_{p+1}. The sequences of No. 16 and the corresponding irreducible chains shall be designated by $\{C_p^{hai}\}$, C_{p+1}^{ha}, so that $C_p^{hai} = C_{p+1}^{ha} \cdot \Phi^i$.

THEOREM II. *$\mathfrak{M}$ possesses a base mod $\mathfrak{N}$, which is in general infinite, and whose elements are irreducible projection-chains.*

Given any particular sequence $\{C_p{}^i\}$, $C_p{}^i \subset \mathfrak{M}^i$, there exists an h, its *index* such that $C_p{}^i \subset \mathfrak{N}^i$ for every $i < h$ but not for $i = h$. Given $C_{p+1} \subset \mathfrak{M}$, we shall call *index* of C_{p+1} the index of the sequence $\{C_{p+1} \cdot \Phi^i\}$. The $(p+1)$-chains whose index $\geqq h$ form a submodulus $\mathfrak{M}_h$ of $\mathfrak{M}$, and we have $\mathfrak{M}_{h+1} \subset \mathfrak{M}_h \subset \mathfrak{M}$. If $C_{p+1} \subset \mathfrak{M}$ we have

$$C_{p+1} \cdot \Phi^h = \Sigma\, t_{ha} C_p^{hai} \quad \text{mod } \mathfrak{N}^h,$$
$$C'_{p+1} = C_{p+1} - \Sigma\, t_{ha} C^{ha}{}_{p+1} \subset \mathfrak{M}_{h+1}.$$

We can treat similarly C'_{p+1}, etc. Ultimately we thus obtain a chain

$$D_{p+1} = \sum_{i \geqq h} t_{ia} C^{ia}{}_{p+1}$$

such that the index of $C_{p+1} - D_{p+1}$ exceeds any positive number. Therefore $C_{p+1} - D_{p+1} \subset \mathfrak{N}$. It is also clear that no element of $\{\{C^{ha}{}_{p+1}\}\}$ can be expressed in terms of those of same or higher index. Therefore $\{\{C^{ha}{}_{p+1}\}\}$ is a base whose elements are irreducible projection-chains.

18. THEOREM III. *There exists a base for the p-cycles mod A, A closed, whose elements are irreducible projection-cycles mod A.*

Let $L^*{}_A$ $(= L^*)$ be the generalized fundamental complex of A. Take for $\mathfrak{M}^h$ the set of all p-cycles of Φ^h mod $\Phi^h \cdot L^*$ such that if Δ_p is any one of them, every Φ^i, $i \geqq h$, contains a chain Δ'_p whose projection on $\Phi^h \approx \Delta_p$ mod $\Phi^h \cdot L^*$ on Φ^h. If Γ_{p+1} is any cycle of K mod $(L^* + \Phi)$ then $\Gamma_{p+1} \cdot \Phi^h \subset \mathfrak{M}^h$ for every h. For take $i > h$, and let $\Gamma_{p+1} \cdot \Phi^i$ be projected onto $\Delta^*{}_p$ of Φ^h. We have

$$\Gamma_{p+1}(\bar{N}^h - N^i) \rightarrow \Gamma_{p+1} \cdot \Phi^h - \Gamma_{p+1} \cdot \Phi^i \text{ mod } L^*.$$

Moreover if D_{p+1} is the deformation-chain corresponding to the projection

$$D_{p+1} \rightarrow \Delta^*{}_p - \Gamma_{p+1} \cdot \Phi^i \text{ mod } L^*.$$

Therefore

$$\bar{N}^h - N^i \supset C_{p+1} \rightarrow \Gamma_{p+1} \cdot \Phi^h - \Delta^*{}_p \text{ mod } L^*.$$

The chain C_{p+1} is finite and by sliding it along the projecting lines we can reduce it to a chain on Φ^h without modifying its boundary mod L^*. Therefore

$$\Gamma_{p+1} \cdot \Phi^h - \Delta^*_p \approx 0 \text{ mod } \Phi^h \cdot L^* \text{ on } \Phi^h; \ \Gamma_{p+1} \cdot \Phi^h \subset \mathfrak{M}^h.$$

The modulus $\mathfrak{M}^h$ also contains all the bounding chains mod $\Phi^h \cdot L^*$ of Φ^h. For E_{p+1} of Φ^h is the projection of some E'_{p+1} of Φ^i; hence $F(E_{p+1}) \subset \mathfrak{M}^h$, and likewise $F(C_{p+1})$ mod $\Phi^h \cdot L^*$ is in $\mathfrak{M}^h$. These bounding cycles form a submodulus $\mathfrak{N}^h$ of $\mathfrak{M}^h$ and the moduli $\mathfrak{M}^h, \mathfrak{N}^h$ are related as in Theorem I. By what we have shown the corresponding moduli $\mathfrak{M}, \mathfrak{N}$ of $(p+1)$-subchains of K are respectively those of the cycles of K mod $(L^* + \Phi)$ and of the bounding cycles of K mod $(L^* + \Phi)$. The required theorem is then a direct consequence of Theorem II.

THEOREM IV. *Any chain c_p is homologous on itself to a projection-cycle mod its own boundary.*

Let $|c_p| = B$, $|F(c_p)| = A$. By Theorem III there is a base $c_p^1, c_p^2, \cdots$, for the p-cycles of B mod A whose elements are irreducible projection-chains which are projection-cycles mod A. Moreover, referring to No. 17, the index $h(\alpha)$ of c_p^α increases indefinitely with α. We have then

$$c_p \approx c'_p = \Sigma\, t_\alpha \cdot c_p^\alpha \text{ mod } A \text{ on } B. \tag{18.1}$$

Let C^α_{p+1} be the projection-chain of K which represents c_p. Among the chains C^α_{p+1} only a finite number have a trace $\neq 0$ on any Φ^i, and each of these chains satisfies the condition for projection-chains. Hence any linear combination of them, and in particular the representative of c'_p, satisfies the same condition. Therefore c'_p is a projection-chain, and (18.1) proves our theorem.

THEOREM V. *If a projection-cycle mod A is ≈ 0 mod A it bounds a projection-chain mod A.*

Let Γ_{p+1} be a projection-cycle mod $(L^* + \Phi)$ representing the projection-cycle mod A, γ_p. Take for $\mathfrak{M}^h$ the modulus of all $(p+1)$-chains on Φ^h whose boundary mod $L^* \cdot \Phi^h$ is a multiple of $\Gamma_{p+1} \cdot \Phi^h$ and for $\mathfrak{N}^h$ the $(p+1)$-cycles of Φ^h mod $L^* \cdot \Phi^h$. Let C_{p+1} be the projection of any element of $\mathfrak{M}^i - \mathfrak{N}^i$ on Φ^h and let $C'_{p+1} \subset \mathfrak{M}^h - \mathfrak{N}^h$. Then

$$tC_{p+1} - C'_{p+1} \to 0 \text{ mod } L^* \cdot \Phi^h, \ t \neq 0.$$

Hence the $(p+1)$-chain at the left is in $\mathfrak{N}^h$. Hence $\mathfrak{M}^h, \mathfrak{N}^h$ are related in the proper way, and by Theorem I there is a projection-chain C_{p+2} such that

$$t_h(C_{p+2} \cdot \Phi^h) \to \Gamma_{p+1} \cdot \Phi^h \text{ mod } L^* \cdot \Phi^h.$$

Since $\{F(C_{p+2} \cdot \Phi^h)\}$ and $\{\Gamma_{p+1} \cdot \Phi^h\}$ are both projection-sequences (up to a chain

of $L^* \cdot \Phi^h$) t_h has a value t independent of h, and $tC_{p+2} \to \Gamma_{p+1}$ mod $(L^* + \Phi)$. Therefore tC_{p+2} represents a projection c_{p+1} of $\mathcal{R}$ which $\to \gamma_p$ mod A.

19. *Connectedness and circuits.* Let again A, L^*_A be a closed set and its generalized fundamental complex, and let x, y be two points of $\mathcal{R} - A$. If $\{\sigma^i\}$ is a projection-sequence for x, any vertex A^i of σ^i is the projection on Φ^i of a vertex A^{i+1} of σ^{i+1}. Therefore x has a projection-sequence $\{A^i\}$ consisting of vertices of the Φ's, and there is a similar sequence $\{B^i\}$ for y.

Now the N S C in order that $\mathcal{R}$ be not disconnected by A is that for every i above a certain value there exist a sequence $U^{i\beta_1}, \cdots, U^{i\beta_r}$ of the sets of the covering Σ^i, in which any two consecutive sets intersect and $U^{i\beta_1} = A^i$, $U^{i\beta_r} = B^i$. This condition is equivalent to $A^i \approx B^i$ on Φ^i for i sufficiently high, and hence for every i. Since $\{A^i\}$, $\{B^i\}$ are projection-sequences they are the traces of projection-cycles Γ_1, Γ'_1 homologous on $K - L^*_A$ mod Φ and representing respectively x, y. By means of Theorem IV and V we find immediately that the above condition is equivalent to $x \approx y$ on $\mathcal{R} - A$. *Therefore the N S C in order that $\mathcal{R} - A$ be connected is that* $R_0(\mathcal{R} - A) = 1$.

Another N S C f,or the connectedness of $\mathcal{R} - A$ is that *the open complexes* $(K - L^*_A) \cdot \Phi^i = \Psi^i$ be all connected. For if they are connected we always have $A^i \approx B^i$ on Ψ^i whatever x, y and hence $x \approx y$ on $\mathcal{R} - A$ so that $\mathcal{R} - A$ is connected. Conversely if $\mathcal{R} - A$ is connected we always have $A^i \approx B^i$ on Ψ^i. But any two vertices A^i, B^i of Ψ^i belong to two projection-sequences $\{A^i\}$, $\{B^i\}$; hence any two vertices of Ψ^i are homologous on Ψ^i; therefore $R_0(\Psi^i) = 1$, and Ψ^i is connected.

20. THEOREM VI. *An n-circuit $\mathcal{R} - A$ is connected and n-dimensional at all points.*

If $\mathcal{R} - A$ is disconnected every Ψ^i, for i above a certain value, is the sum of two open complexes without common cells. As a consequence, by suppressing a suitable finite part of K, we shall have a new K such that $K - L^*_A = K' + K''$, where K' and K'' are open complexes without common cells. Let γ_n be the fundamental n-cycle mod A of the circuit and let Γ_{n+1} be its representative cycle mod $(L^*_A + \Phi)$. The chains

$$\Gamma'_{n+1} = K' \cdot \Gamma_{n+1}, \qquad \Gamma''_{n+1} = K'' \cdot \Gamma_{n+1}$$

are similar cycles which represent cycles mod A, γ'_n and γ''_n, such that

$$\gamma_n = \gamma'_n + \gamma''_n, \qquad |\gamma'_n| \cdot |\gamma''_n| \subset A.$$

Any point x of γ'_n has a neighborhood $U \subset \mathcal{R} - A - |\gamma''_n|$. Hence $B = \mathcal{R} - U$ is a closed set $\supset A$ and also γ''_n, so that $R_n(B; A) \neq 0$, which contradicts one of the circuit conditions. Therefore $\mathcal{R} - A$ is connected.

Regarding the dimension of $\mathfrak{R} - A$, let $x \subset \mathfrak{R} - A$, $\dim_x \mathfrak{R} = p < n$. We can find an open set $U \supset x$ such that $U \subset \mathfrak{R} - A$, $\dim F(U) \leqq p - 1$. It follows that if we suppress a suitable finite part of K, the new $K - L^*_A$ shall be disconnected into two subcomplexes K', K'' by a projection-complex L which is a fundamental complex for $F(U)$ and whose dimension is therefore $\leqq p$ (No. 7). Since $p < n$ neither K' nor K'' will have $(n+1)$-cells with n-faces on L. Hence $K' \cdot \Gamma_{n+1}$ and $K'' \cdot \Gamma_{n+1}$ are separately cycles mod $(L^*_A + \Phi)$ determining cycles mod A, γ'_n and γ''_n whose intersection $\subset A$, and we have the same contradiction as before.

21. *Extension to locally compact spaces.* Practically all our results may be extended to a locally compact separable space $\mathfrak{R}$. It is known that such a space is metric and that it can be mapped topologically on a compact space $\mathfrak{R}^*$ with a point x^* removed. That is to say $\mathfrak{R}$ can be identified with $\mathfrak{R}^* - x^*$. Moreover, topologically speaking $\mathfrak{R}^*$ is unique.† If U^* is a neighborhood of x^* on $\mathfrak{R}^*$, $U = U^* - x^*$ is an open set of $\mathfrak{R}$ and $F(U^*) = F(U)$. There exists then another such set $V^* \Subset U^*$, and if $V = V^* - x^*$, we have also on $\mathfrak{R} : V \Subset U$. Since $\mathfrak{R}$ is n-dimensional we can find an open set W of $\mathfrak{R}$ such that $V \Subset W \Subset U$, $\dim F(W) < n$. Therefore if $W^* = W + x^*$, we have $\dim F(W^*) < n$, $x^* \subset W^* \Subset U^*$. In other words given any neighborhood U^* of x^* there is another $W^* \Subset U^*$ whose boundary is of dimension $< n$. This shows that $\dim_{x^*} \mathfrak{R}^* \leqq n$. Any point $x \neq x^*$, has relatively to $\mathfrak{R}^*$ a neighborhood which is also a neighborhood relatively to $\mathfrak{R}$ and hence $\dim_x \mathfrak{R}^* = \dim_x \mathfrak{R}$, $\dim \mathfrak{R}^* = \dim \mathfrak{R} = n$.

Let K^* be a fundamental complex of $\mathfrak{R}^*$, and let $\{\sigma^i\}$ be a fundamental sequence for x^* and L^* the sum of the sequence and its joining cells. Then $K = K^* - L^*$ is an open complex which we may consider as a fundamental complex for $\mathfrak{R}$. We now have two types of cycles to consider for $\mathfrak{R}$: (a) the finite p-cycles; they correspond to the $(p+1)$-cycles of $K = K^* - L^*$ mod Φ; (b) the infinite p-cycles of $\mathfrak{R}$; they are represented by the $(p+1)$-cycles of K^* mod $(L^* + \Phi)$. Both types have essentially the same properties as the cycles previously considered. It is also a simple matter to show that they are topological elements of $\mathfrak{R}$ itself, independent of the mode of turning $\mathfrak{R}$ into a compact space $\mathfrak{R}^*$. Let us state in passing that $\mathfrak{R}$ will be called an *open n-circuit* whenever it behaves in the same manner regarding the infinite n-cycles as previously regarding the finite (ordinary) n-cycles.

It is to be observed that the space $\mathfrak{R}$, or rather the set $\mathfrak{R}$, may actually

† Urysohn and Alexandroff, *Mémoire sur les spaces topologiques compacts*, Amsterdam Academy, Verhandeligen, Deel XIV, No. 1, 1929.

be given in the form $\mathfrak{R}^* - A$, where $\mathfrak{R}$ is compact, metric, but not necessarily n-dimensional, and A is a closed subset of $\mathfrak{R}^*$ that may consist of more than one point. L^* is then merely a fundamental complex for A, but otherwise the rest is as before. We can pass to the case where A is a single point by applying to $\mathfrak{R}^*$ a continuous single-valued transformation, homeomorphic over $\mathfrak{R}^* - A$, and reducing A to a single point. Concurrently we replace the subcomplex $L^* \cdot \Phi^i$ by a single point and L^* by a single projection line. It is clear that this does not affect the cycles which we have introduced above nor their homologies.

PART II. The Generalized Manifold.

22. *Definition.* A generalized n-manifold M_n or M is a locally compact separable n-space with the following properties whose topological character is obvious:

I. M is the sum of a countable aggregate of disjoined n-circuits.

II. The Betti-number $R_n(M, M - x) = 1$ for every $x \epsilon M$.

III. M is locally connected.

IV. Given any closed q-set F on M_n and any open set U there is an open set $V \subset U$, such that every chain c_p, $p < n - q$, on V, whose boundary does not meet F, is deformable over U, without moving its boundary, into a chain c'_p which does not meet F.

If the circuits in I are absolute we call M an *absolute* manifold, otherwise an *open* manifold. If the circuits are all *orientable* M is orientable, otherwise it is *non-orientable.*

Interpretation of the manifold conditions. Condition I requires no comment. Regarding II we may consider, with van Kampen, as p-cycle of a point x a c_p whose boundary $\not\supset x$, i. e. a cycle mod $M - x$ in the sense of *Topology*, the homologies being of the type $\approx M - x$: $c_p \approx 0 \bmod M - x$ means that there is a c_{p+1} such that $F(c_{p+1}) - c_p \not\supset x$. The corresponding Betti-number is the number designated by $R_p(M, M - x)$. Since the fundamental γ_n of M is itself a cycle mod $M - x$ whatever $x \epsilon M$, II signifies that if c_n is any chain whose boundary $\not\supset x$, a certain $c_n - t\gamma_n \not\supset x$.

The local connectedness in III is the so-called local zero-connectedness of *Topology*, p. 90. It means explicitly that for every U there is a $V \subset U$ such that any two points of V are on a connected subset of U.

When we have an absolute M conditions III and IV are equivalent to:

III′. There exists, for every $\epsilon > 0$, a number $\delta(\epsilon) < \epsilon$, such that any two points not farther apart than δ, are on a connected set of diameter $< \epsilon$.

IV′. For every F and ϵ there exists a number $\eta(\epsilon) < \epsilon$, such that any c_p of diameter $< \eta$ whose boundary does not meet F, where $\dim F = q < n - p$, is ϵ-deformable without displacing its boundary, into a chain which does not meet F.

For $0 < p < n$, $F = 0$, this becomes a weak type of local q-connectedness with the p-cell and $(p-1)$-sphere of *Topology* replaced by a p-chain and $(p-1)$-cycle.

Conditions II, III, IV are purely local and serve to characterize the homogeneity properties of M. Condition I on the contrary refers to the whole manifold and serves also to separate the different types.

23. *The Kronecker-index. Definition.* Taking, merely for convenience, an absolute M_n, let c_p, c_{n-p} be two chains on M_n which do not intersect one another's boundaries:

$$(23.1) \qquad | c_p | \cdot | F(c_{n-p}) | + | F(c_p) | \cdot | c_{n-p} | = 0.$$

Their Kronecker-index is to be a number $(c_p \cdot c_{n-p})$ such that:

(a) $(c_p \cdot c_{n-p}) = 0$ when the chains do not meet.

(b) The index when defined is a bilinear function of the two chains.

(c) If the boundaries of c_p, c_{n-p+1} do not intersect then

$$(23.2) \qquad (c_p \cdot F(c_{n-p+1})) = (-1)^p (F(c_p) \cdot c_{n-p+1}).$$

(d) If $x \epsilon M$ and γ_n is the fundamental n-cycle of M then

$$(23.3) \qquad (x \cdot \gamma_n) = 1.$$

We shall show that there exists a unique index which is a topological invariant of the two chains and which satisfies conditions (a), $\cdots$, (d), and has the following additional properties:

(e) If c_{p+1} and $F(c_{n-p})$ do not meet

$$(23.4) \qquad (F(c_{p+1}) \cdot c_{n-p}) = 0,$$

and similarly with p and $n-p$ interchanged.

(f) The chains being as in the definition:

$$(23.5) \qquad (c_p \cdot c_{n-p}) = (-1)^{p(n+1)} \cdot (c_{n-p} \cdot c_p).$$

The existence proof as well as properties (e), (f), will be established by induction.

In the theory of the index for combinatorial manifolds (*Topology*, Ch. IV), (a), (b), (d) enter more or less in the definition, while (c) is proved

explicitly. It is in fact essentially formula (20) *loc. cit.*, which plays an all important part there and is a direct consequence of the fundamental boundary relation (18) for intersections of chains. On the contrary here the same relation serves directly to define recurrently the index without passing through intersections of dimension > 0. This is substantially in accord with the definitions suggested *loc. cit.*, p. 216. See also H. A. Newman's recent paper *Cambridge Philosophical Transactions*, Vol. 27 (1931), pp. 491-501.

24. The Kronecker-index $(c_0 \cdot \gamma_n)$, where c_0 is a projection-chain and γ_n the fundamental n-cycle, is readily treated. In the first place if C_0 is a finite subchain of a complex,

(24.1) $$C_0 = \Sigma\, t_i E_0{}^i,$$

we define its Kronecker-index as in *Topology*, p. 169, by

(24.2) $$(C_0) = \Sigma\, t_i,$$

and we recall that $C_0 \approx 0$ implies that $(C_0) = 0$. If $\mathfrak{K}$ is an orientable and oriented combinatorial manifold we have $(C_0) = (C_0 \cdot \mathfrak{K})$.

Let now c_0 be a projection-chain of M (projection-zero-cycle), with Γ_1 as its representative projection-cycle mod Φ on K. Since $\tau_i \Gamma_1 \cdot \Phi^{i+1} = \Gamma_1 \cdot \Phi^i$, we have $(\Gamma_1 \cdot \Phi^{i+1}) = (\Gamma_1 \cdot \Phi^i)$. This index is therefore independent of i and its value is by definition (c_0).

We shall show below, that c_0 can be ϵ-deformed whatever ϵ, into a chain consisting of a finite number of points $x_0, \cdots, x_r$, so that

(24.3) $$c_0 \approx \Sigma\, s_j x_j.$$

Since M is connected, if x is any point of M we have $x \approx x_j$, and hence

(24.4) $$c_0 \approx x \,\Sigma\, s_0 = sx.$$

As we have seen (No. 19) x has a projection-sequence $\{A^i\}$ made up of vertices of the Φ's. Owing to (23.4) we have

(24.5) $$\Gamma_1 \cdot \Phi^i \approx sA^i \text{ on } \Phi^i;$$

hence $(\Gamma_1 \cdot \Phi^i) = s(A^i) = s = (c_0)$. Since s is clearly a topological function of c_0 and does not depend in any sense on the fundamental complex K the same holds for (c_0).

If we now define $(c_0 \cdot \gamma_n)$ by the relation

(24.6) $$(c_0 \cdot \gamma_n) = s(x \cdot \gamma_n),$$

we have by the above and (23.3)

(24.7) $$(c_0 \cdot \gamma_n) = (c_0).$$

This disposes of $(c_0 \cdot \gamma_n)$ and shows in particular that it is a topological invariant of c_0.

25. THEOREM VII. (*Deformation theorem*). *Given a projection-chain c_p, a closed q-set F and any ϵ, there exists an ϵ-deformation of c_p into a chain c'_p with an elementary ϵ-decomposition into projection-chains $\{c'_r{}^i\}$, whose zero-chains are isolated points and whose r-chains, $r < n - q$, do not meet F.*

Let c_p be represented by C_{p+1} of K and let $E_q{}^{ha}$ be the cells of Φ^h. We decompose C_{p+1} in a sum of r chains:

(25.1) $$C_{p+1} = C^1{}_{p+1} + \cdots + C^r{}_{p+1},$$

where r is the number of cells of Φ^h, and where $C^a{}_{p+1}$ is a subchain of the complex consisting of all the cells of K on the set of all projecting lines that meet E^{ha}. A similar decomposition may then be applied to the chains of $F(C^a{}_{p+1})$, etc., until finally we have an elementary decomposition $\{C^a{}_{q+1}\}$ of C_{p+1} characterized by the property that $C^a{}_{q+1}$ is on the set of cells of K that are on the projecting lines through the points of E^{ha}. The chains $C^a{}_{q+1}$ like C_p itself, are projection-chains. There results an elementary decomposition $\{c_q{}^a\}$ of c_p into projection-chains associated with each Φ^h.

Let us now observe that the points of M in whose projection-sequences $\{\sigma^i\}$ the term σ^k is E^{ka} or a face of it, are the points of sets $U^{k\beta}$ with a common point. Their sum is an open set V^{ka} whose diameter $< 2\epsilon_k$, where $\epsilon_k =$ mesh Σ^k. Let η_k be the characteristic-constant of the f. c. o. s. $\{V^{ka}\}$ and take h so high that $\epsilon_h < \frac{1}{2}\delta(\eta_k)$, where δ is the same as in No. 22, III'. As a consequence any two points x, y on a set $V^{h\gamma}$ will be on a connected subset of some $\bar{V}^{ka} \supset V^{h\gamma}$. By No. 19, $x \approx y$ on $\bar{V}^{ka}$. It follows that if $\{\sigma^i\}$, $\{\sigma'^i\}$ are representative sequences for x, y then for $i > h$, σ^i and σ'^i can be joined on Φ^i by a polygonal arc λ (sum of vertices and one-cells of Φ^i) whose projections on Φ^k is on E^{ka}. Hence any two vertices of the subcomplex of Φ^i projected onto E^{ha} can be joined in the above manner by a polygonal arc on Φ^i. For both belong to a pair of simplexes such as σ^i, σ'^i.

26. Henceforth h, k are to be kept fixed. Since M is an n-circuit it is n-dimensional at all points (Theorem VI). Since $q < n$ there exists then on every open set a point not on F. Choose such a point x_a on V^{ha}, and let $\{A^{ai}\}$ be a projection-sequence of vertices for the point x_a.

Consider one of the zero-chains $c_0{}^a$, of the decomposition of c_p. It is defined by means of a certain projection-chain $C_1{}^a$ of K so that $\{C_1{}^a \cdot \Phi^i\}$ is a

projection-sequence of zero-chains. From the above follows immediately that when i exceeds a certain value we can find a one-chain D_1^{ai} of Φ^i whose projection is on E^{ka} and such that in addition

$$(26.1) \qquad D_1^{ai} \rightarrow (C_1^a \cdot \Phi^i)A^{ai} - C_1^a \cdot \Phi^i = (c_0^a)A^{ai} - C_1^a \cdot \Phi^i.$$

By Theorem V we may choose for D_1^{ai} a projection-sequence which determines a projection-chain D_2^a of K, and finally a projection-chain d_1^a of $\mathfrak{R}$. As a consequence of (26.1) we have then

$$(26.2) \qquad d_1^a \rightarrow (c_0^a)x_a - c_0^a.$$

We have thus displaced the zero-chains of the elementary decomposition of c_p into points $\not\subset F$. The displacement and the diameters of the chains of the decomposition may be made as small as we please. From this point on and taking account of Theorem V and condition IV of No. 22 for an M_n, the proof of the required deformation theorems proceeds as in *Topology*, p. 93. The only modifications are that singular cells and thei rsingular boundary spheres are replaced by the elementary chains c_q^i and their boundaries.

27. As a first application let $x \subset c_p - F(c_p) \neq 0$, where c_p is otherwise arbitrary. The chain is homologous on itself mod its boundary to a projection-chain c'_p so that $x \not\subset F(c'_p)$. By Theorem VII c_p can be ϵ-deformed whatever ϵ into a projection-chain $c_p'' \not\supset x$. This implies that $c_p \approx 0$ on $M - x$, and also that for every $\xi > 0$ there is an $\eta(\xi)$ such that if x is a point of c_p farther than ξ from $F(c_p)$, then $c_p \approx c'_p$, where c'_p is at a distance $\geqq \eta$ from x.

Referring to No. 22 we have by what precedes,

$$(27.1) \qquad R_p(M - x) = \delta_{np},$$

where δ_{np} is the Kronecker delta ($= 1$ for $p = n$, $= 0$ for $p \neq n$).

Let us now observe that if we apply the construction of Nos. 24, 25 to a γ_p whose diameter is sufficiently small, we may choose all the chains c'^i_q coincident with a single point x. As a consequence if $p > 0$, $c'_p = 0$, and by (11.1) the deformation chain

$$(27.2) \qquad \mathfrak{D}\gamma_p \rightarrow -\gamma_p,$$

Therefore for every open set U there is another $V \subset U$ such that every γ_p, $0 < p < n$, on V is ≈ 0 on U.

The preceding statement is valid for any manifold. For an absolute manifold owing to compactness we have: *for every θ there is a $\tau(\theta)$ such that every γ_p, $p < n$, whose diameter $< \tau$ bounds a chain of diameter $< \theta$.* This is

merely another formulation of the weak local p-connectedness property mentioned in No. 22.

28. Let γ_p be one of the irreducible projection-cycles of the base constructed in No. 18 and whose index $i > h$. When we apply the deformation of the preceding numbers with $F = 0$, we find that the deformed cycle $\gamma'_p = 0$, for its chains correspond element for element to the chains of a degenerate simplicial p-cycle. Therefore (27.2) will hold here also and hence $\gamma_p \approx 0$ on M. In particular the base alluded to can only contain a finite number of cycles $\not\approx 0$, namely those whose indices do not exceed a certain value. This proves the important

THEOREM VIII. *The Betti-numbers of an absolute M_n are all finite.*

29. *Determination of the Kronecker-index.* We propose to give a recurrent determination of $(c_p \cdot c_{n-p})$ for two chains which do not intersect one another's boundaries. Taking first $p > 0$ we shall reduce the case in question to the same for $p - 1$ and ultimately to a $(c_0 \cdot c_n)$ where c_0 consists of a finite number of isolated points. This last index shall be treated directly by reduction to the case considered in No. 23. At the same time we shall show that the index has all the properties expected. We assume then first that this holds already for $p - 1$, extend it to p, then take up the case $p = 0$ at the end.

Our first move is to replace c_p, c_{n-p} by projection-chains, homologous respectively to c_p, c_{n-p} on $|c_p|$, $|c_{n-p}|$ mod their boundaries (Theorem IV). To simplify matters we continue to denote the new chains by c_p, c_{n-p}. We merely recall that after the reductions the new sets $|c_p|$, $|c_{n-p}|$, $|F(c_p)|$, $|F(c_{n-p})|$ are subsets of the old. As a consequence in what follows, $|c_p|$ for example, may designate indifferently the new or the old set $|c_p|$.

Let now ξ be the least of the two positive numbers $d(c_p, F(c_{n-p}))$, $d(F(c_p), c_{n-p})$. Since every point x of $|c_p|$ is at least as far as ξ from $F(c_{n-p})$, x has a neighborhood V such that $c_{n-p} \approx 0 \bmod M - V$. Since $|c_p|$ is self-compact it can be covered with a finite number of neighborhoods $V^1, \cdots, V^r$, such that $c_{n-p} \approx 0 \bmod M - V^j$. That is to say there exists a projection-chain

$$(29.1) \qquad c^j_{n-p+1} \to c_{n-p} \bmod M - V^j.$$

The sets V^j form a f. c. o. s. for $|c_p|$ and there is an analogue ζ of the characteristic constant for that covering: every point x of $|c_p|$ will be on some V^j such that $d(x, M - V^j) > \zeta$.

We shall now choose a certain fixed $\epsilon > 0$, and the determination of the index will depend upon that ϵ. We shall endeavor to show that the index remains the same for all ϵ's sufficiently small, and so we shall not hesitate to

take this ϵ arbitrarily small. In particular we shall require that $\epsilon > \frac{1}{4}\xi, \frac{1}{4}\zeta$ or $\tau(\frac{1}{2}\zeta)$ where τ is the same function as in No. 27.

By Theorem VII we may ϵ-deform c_p into a projection-chain c'_p with an elementary ϵ-decomposition $\{c'^i_q\}$ whose elements (all projection-chains also) of dimension $q < p$ do not meet c_{n-p} and with

$$c'_p = \Sigma\, c'^i_p. \tag{29.2}$$

The chains c'^i_p are in finite number and their boundaries do not meet c_{n-p}. We shall set by definition $(c_p \cdot c_{n-p}) = (c'_p \cdot c_{n-p})$, and hence, if (23b) is to hold,

$$(c_p \cdot c_{n-p}) = (c'_p \cdot c_{n-p}) = \Sigma(c'^i_p \cdot c_{n-p}), \tag{29.3}$$

where in the sum we preserve only the useful terms, namely those corresponding to chains c'^i_p which meet c_{n-p}. The problem is to determine the indices in that sum.

Since the points of c'^i_p are not farther than $\epsilon < \frac{1}{4}\xi$ from c_p, and since its diameter $< \frac{1}{4}\zeta$, c'^i_p is on at least one set V^j and farther than $\frac{1}{2}\zeta$ from the corresponding $M - V^j$. Choose one of the sets V^j of this nature and relabel that V and its c_{n-p+1} entering in (29.1), respectively V'^i, c'^i_{n-p+1}. The situation being as described we have $F(c'^i_{n-p+1}) = c_{n-p} + c'_{n-p}$, where c'_{n-p} does not meet c'^i_p. Hence, if the basic laws (23a c) for the index are to hold, we must have

$$(c'^i_p \cdot c_{n-p}) = (-1)^p\,(F(c'^i_p) \cdot c'_{n-p+1}), \tag{29.4}$$

and hence by (23 b)

$$(c_p \cdot c_{n-p}) = (-1)^p\, \Sigma\, (F(c'^i_p) \cdot c'^i_{n-p+1}). \tag{29.5}$$

The right hand side is known under the hypothesis of the induction, hence (29.5) determines a value for $(c_p \cdot c_{n-p})$. It remains to be shown that the index thus obtained behaves as expected.

30. We first observe that the index is a linear function of c_p. That is to say if the preceding method has yielded $(c_p \cdot c_{n-p})$ and $(c'_p \cdot c_{n-p})$ then it also yields

$$(tc_p + t'c'_p \cdot c_{n-p}) = t(c_p \cdot c_{n-p}) + t'(c'_p \cdot c_{n-p}). \tag{20.1}$$

We notice also that if c_p and c_{n-p} do not meet and if we take ϵ less than half their distance apart, the value computed for their index is zero, and hence accords with (23 a) independently of the variable elements entering in its determination.

Let us replace V'^i by any other V, say V''^i, behaving in the same manner relatively to c'^i_p and let c''^i_{n-p+1} be the corresponding c_{n-p+1}. To show that

substituting V''^i for V'^i has not modified the index we must prove that if we substitute c''^i_{n-p+1} for c'^i_{n-p+1} in (29.5) the index remains the same. Since (23 b) holds for $p-1$ this merely requires that we prove

$$(30.2) \qquad \Sigma\,(F(c'^i_p)\cdot c'^i_{n-p+1}-c''^i_{n-p+1})=0.$$

Let $W=V'^i\cdot V''^i$. This open set $\supset c'^i_p$ and $d(c'^i_p, M-W)>\frac{1}{2}\zeta$. Moreover both chains c'^i_{n-p+1}, $c''^i_{n-p+1}\rightarrow c_{n-p} \bmod M-W$. Hence

$$(30.3) \qquad c'^i_{n-p+1}-c''^i_{n-p+1}\rightarrow 0 \bmod M-W.$$

On the other hand since diam $c'^i_p<\tau(\frac{1}{2}\zeta)$ and since, by hypothesis, (24 e) holds for $p-1$ in place of p, we have

$$(30.4) \qquad (F(c'^i_p)\cdot c'^i_{n-p+1}-c''^i_{n-p+1})=0,$$

from which the required relation (30.2) follows. If $V'^i=V''^i$ but c'^i_{n-p+1} is replaced by c''^i_{n-p+1} the same reasoning holds. Therefore a modification in the chains c_{n-p+1} likewise leaves the index unaltered.

31. Let us now show that (24 e) holds: if

$$(31.1) \qquad M-F(c_{n-p})\supset c_{p+1}\rightarrow c_p$$

then we have

$$(31.2) \qquad (c_p\cdot c_{n-p})=0.$$

Take a $U\supset c_{p+1}$ and $\Subset M-F(c_{n-p})$, then apply Theorem V with $\bar{U}$ as the basic space. As a consequence we find that we may assume that c_{p+1} is a projection-chain. By Theorem VII c_{p+1} is ϵ-deformable into c'_{p+1} with an ϵ-decomposition $\{c'^i_q\}$ whose chains of dimension $<p$ do not meet c_{n-p}. It is to be observed that the construction of the deformed chains is such that the chains c'^i_r depend solely on those of dimension $<r$. Hence c_p *is* thus deformed into any c'_p serving to calculate its index in accordance with No. 29. We shall have as the new p- and $(p+1)$-chains

$$(31.3) \qquad c'_{p+1}=\Sigma\, c'^i_{p+1}, \qquad c'_p=\Sigma\, F(c'^i_{p+1}),$$

and therefore

$$(31.4) \qquad (c_p\cdot c_{n-p})=\Sigma\,(F(c'^i_{p+1})\cdot c_{n-p}).$$

If ϵ is small enough the chains c'^j_{p-1} on $F(c'^i_{p+1})$ will meet a single chain c_{n-p+1} that we may call as before c'^i_{n-p+1}. By (29.5) and No. 30

$$(31.5) \qquad (F(c'^i_{p+1})\cdot c_{n-p})=(-1)^p\,(F(F(c'^i_{p+1}))\cdot c'^i_{n-p+1})\equiv 0,$$

since $F(F)\equiv 0$, and from this follows (31.2).

As an application suppose that we have obtained, always by means of K, two different ϵ-deformations $\mathfrak{D}'$, $\mathfrak{D}''$ of c_p into c'_p and c''_p serving to calculate the index $(c_p\cdot c_{n-p})$. We have

(31. 6) $\mathfrak{D}'c_p \to c'_p - c_p - \mathfrak{D}'F(c_p), \qquad \mathfrak{D}''c_p \to c''_p - c_p - \mathfrak{D}''F(c_p),$

(31. 7) $$\mathfrak{D}'c_p - \mathfrak{D}''c_p \to (c'_p - c''_p) - \cdots$$

where, under the limitations upon ϵ, the chain omitted does not meet c_{n-p}. Hence

(31. 8) $$(c'_p - c''_p \cdot c_{n-p}) = 0,$$

whatever the procedure chosen to compute the index. Take as the deformation the process which consists merely in replacing c'_p by the decomposition associated with the deformation $\mathfrak{D}'$, and similarly for c''_p and $\mathfrak{D}''$. As a consequence the index (31. 8) becomes merely the difference of the values of the index $(c_p \cdot c_{n-p})$ as computed by means of the two deformations. Therefore these two values are the same. In other words $(c_p \cdot c_{n-p})$ *is independent of the ϵ-deformations used in computing it.*

32. We have already shown that our index possesses properties (23 a) and part of (23 b e). We still have to show that when $p > 0$ properties (23 b c e f) hold.

Since we have established the linearity of the index in c_p, (23 b) will be established if we show that the index is also linear in c_{n-p}. Consider two projection-chains c_{n-p}, c'_{n-p} and let them not meet $F(c_p)$, (c_p a projection-chain), nor let c_p meet their boundaries. Let the ϵ-deformation of c_p into c'_p be so carried out that the q-chains c'^i_q, $q < p$, of the decomposition of c'_p meet neither c_{n-p} nor c'_{n-p}. Then it follows at once from the definition of the index by (29. 5) that

(32. 1) $$\begin{aligned}(c_p \cdot tc_{n-p} + t'c'_{n-p}) &= (c'_p \cdot tc_{n-p} + t'c'_{n-p}) \\ &= t(c'_p \cdot c_{n-p}) + t'(c'_p \cdot c'_{n-p}) = t(c_p \cdot c_{n-p}) + t'(c_p \cdot c'_{n-p}),\end{aligned}$$

which proves the required linearity and hence also that property (b) holds completely.

Consider now property (c): if c_p, c_{n-p+1} have non-intersecting boundaries then (23. 2) holds. Here we may take in (29. 5) every $c'^i_{n-p+1} = c_{n-p+1}$, which yields

(32. 2) $$(c_p \cdot F(c_{n-p+1})) = (-1)^p \Sigma\, (F(c'^i_p) \cdot c_{n-p+1}).$$

In the summation in (29. 5) only certain chains c'^i_p whose sum is c'^i_p were preserved, namely those which met c_{n-p}. As we have just shown if c'^i_p does not meet c_{n-p} the corresponding contribution of its boundary to the sum in (32. 1) is zero; hence the summation may now be extended to all the chains c'^i_p. By the linearity of the index for $p - 1$, $n - p + 1$ we have:

(32. 3) $$\begin{aligned}\Sigma(F(c'^i_p) \cdot c_{n-p+1}) &= (F(\Sigma\, c'^i_p) \cdot c_{n-p+1}) \\ &= (F(c'_p) \cdot c_{n-p+1}) = (F(c_p) \cdot c_{n-p+1}).\end{aligned}$$

This relation together with (32.2) yields (23.2) and proves that property (c) holds.

From (a) and (c) follows that if $F(c_p)$ and c_{n-p+1} do not meet then

$$(c_p \cdot F(c_{n-p+1})) = 0. \tag{32.4}$$

This is the analogue of (23.4) with p and $n-p$ interchanged, and together with the result of No. 31 it embodies the proof of property (e).

We postpone the proof of property (f) till later.

33. We shall now consider the case $p=0$, i. e. the index $(c_0 \cdot c_n)$, where as before $c_0 \subset M - F(c_n)$. As in No. 24, we have here for an ϵ sufficiently small the analogue of (24.3):

$$c_0 \approx \Sigma\, s_j x_j \text{ on } M - F(c_n). \tag{33.1}$$

Now for x_j we have by No. 22, condition II,

$$c_n \approx t_j \gamma_n \bmod M - x_j, \tag{33.2}$$

and we shall set

$$(c_0 \cdot c_n) = \Sigma\, s_j t_j (x_j \cdot \gamma_n) = \Sigma\, s_j t_j. \tag{33.3}$$

Properties (a), (b) are at once verified for this index and we only have to prove (e), (f). Here also (f) shall be treated later.

The proof of (e) consists of two parts:

(a) if $M - F(c_n) \supset c_1 \rightarrow c_0$ then $(c_0 \cdot c_n) = 0$. As in No. 31 we may assume that diam $c_1 < \epsilon$ assigned. Now for c_n and any x there is a neighborhood $V \supset x$ such that $c_n - \lambda\gamma_n \approx 0$ on $M - V$. Since M is compact it may be covered with a finite number of such sets V : $V^1, \cdots, V^r$ with $\lambda = \lambda_i$ on V^i. Let us take $\epsilon < \frac{1}{2}\eta$, where η is the characteristic constant of this f. c. o. s., and let the deformations be $< \frac{1}{2}\eta$. We shall take $c_1 \subset V^h$ and farther than $\frac{1}{2}\eta$ from $M - V^h$. Hence if we calculate the index of $c_0 = F(c_1)$ by our method, the corresponding points x_j are all on V^h, and the associated constants t_j all equal to λ_h. Finally since $c_0 \approx 0$, $(c_0) = 0$. Therefore

$$(c_0 \cdot c_n) = (F(c_1) \cdot c_n) = \lambda_h \Sigma\, s_j = \lambda_h (c_0) = 0. \tag{33.4}$$

(b) if $c_n \approx 0$ then $(c_0 \cdot c_n) = 0$. This is evident for $c_n \approx 0$ implies that the representative projection-cycle Γ_{n+1} of c_n is $\equiv 0$ and hence $c_n \equiv 0$. Consequently in the homologies $c_n \approx t\gamma_n \bmod M - x$, we always have $t = 0$, so that $(c_0 \cdot c_n) = 0$.

As in No. 31 the first case considered proves here also that $(c_0 \cdot c_n)$ has a value independent of the particular mode of determining it.

34. Let us return to $(c_p \cdot c_{n-p})$. We have obtained its value by an induction on p in which there appear certain intermediary chains c_{p-i}, c_{n-p+i}, so that we have:

$$
\begin{aligned}
&(c_p \cdot c_{n-p}) = (-1)^p \, (c_{p-1} \cdot c_{n-p+1}), \\
&\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
(34.1) \qquad &(c_{p-i} \cdot c_{n-p+i}) = (-1)^{p-i} (c_{p-i-1} \cdot c_{n-p+i+1}), \\
&\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
&(c_1 \cdot c_{n-1}) = - \, (c_0 \cdot c_n);
\end{aligned}
$$

and hence, in the last analysis,

$$(34.2) \qquad (c_p \cdot c_{n-p}) = (-1)^{p(p+1)/2} (c_0 \cdot c_n) = (-1)^{p(p+1)/2} \cdot \lambda \cdot (x \cdot \gamma_n),$$

the value of λ being the number given by (33.3). Observe that if $p > 0$, the various chains of dimension $p-1, \cdots$, zero, introduced in this determination are chains of elementary decompositions in which the zero-chains consist of isolated points taken with finite multiplicities. Therefore in particular c_0 is of this nature. It follows that the numbers s_j, t_j of No. 33, that serve to compute λ are all finite and so is λ. An immediate consequence is the fact that $(c_p \cdot c_{n-p})$ *is independent of the fundamental complex* K, *and hence the Kronecker-index is a topological invariant.* For if we have any index whatever with the properties (a), $\cdots$, (e) of No. 23, it will satisfy the relations (34.1) and (34.2). Since λ depends solely on certain homologies but not on K our assertion follows.

Now the above has been obtained as a consequence of an induction on p. By means of (23c), explicitly proved for our index, we may carry through a similar induction on $n-p$. This leads to a formula analogous to (34.2)

$$(34.3) \qquad (c_p \cdot c_{n-p}) = (-1)^{[n(n+1)-p(p+1)]/2} \cdot \mu(\gamma_n \cdot x).$$

If we apply the process just stated to $(c_{n-p} \cdot c_p)$ we find that the geometric operations carried out for its determination are the same as those used in determining $(c_p \cdot c_{n-p})$ by our initial procedure (induction on p), and that as a consequence the corresponding μ is λ, both being equal to a certain expression $\Sigma s_j t_j$ appearing in (33.3). For each j the number s_j is the multiplicity of a certain point as constituent of c_0 and t_j the coefficient t in a certain homology (33.2). Therefore

$$(34.4) \qquad (c_{n-p} \cdot c_p) = (-1)^{p(2n-p+1)/2} \lambda \cdot (\gamma_n \cdot x),$$

$$(34.5) \qquad (c_{n-p} \cdot c_p) = (-1)^{p(n+1)} (c_p \cdot c_{n-p}) [(\gamma_n \cdot x)/(x \cdot \gamma_n)].$$

In particular for $p = n$, $c_0 = x$, $c_p = \gamma_n$:

$$(x\cdot\gamma_n)^2 = (\gamma_n\cdot x)^2, \tag{34.6}$$

and hence finally

$$(\gamma_n\cdot x) = \pm(x\cdot\gamma_n). \tag{34.7}$$

We shall prove later that the proper sign to be chosen here is $+$. Assuming this for the present we have from (23 d), $(\gamma_n\cdot x) = 1$ and hence finally

$$(c_{n-p}\cdot c_p) = (-1)^{p(n+1)}(c_p\cdot c_{n-p}) \tag{34.8}$$

which is (23 f). Thus except for a certain choice of sign we have finally established that the Kronecker-index has all the properties required.

35. *Duality properties of the absolute M_n.* In order to obtain the extension of Poincaré's duality relation for the Betti-numbers, all that is now needed is a converse of property (c) of No. 23; if a cycle $\gamma_p \not\approx 0$ there is a γ_{n-p} such that $(\gamma_p\cdot\gamma_{n-p}) \neq 0$. A slightly more general result will now be proved.

Consider first the sequence of the images of the skeleta Φ^i which we still call Φ^i, on an S_{2n+1}, whereby one may map topologically a compact metric n-space, here our absolute M_n, on the space S_{2n+1} † and let us modify the construction as follows: We take an S_{2n+2} referred to coördinates $x_1,\cdots,x_{2n+2}$ and assume that our S_{2n+1} is the one given by $x_1 = 0$, so that M is now mapped onto that space. We then project Φ^i onto the space $x_1 = 1/i$, and replace Φ^i by its projection which we henceforth call Φ^i. The joining cells being inserted as before, if their (linear) spaces happen to have intersections of too high dimension, we may slightly displace the vertices of the Φ's in their $(2n+1)$-spaces so as to remove this untoward circumstance. We now have M and K immersed in a certain S_{2n+2}. We may in fact immerse S_{2n+2} in any S_r, $r \geqq 2n+2$ and together with it also both M and K. We shall choose r such that $r - n$ is even.

Let us surround each Φ^i by a closed polyhedral neighborhood $\mathcal{K}^i$ in S_r, take a subdivision $\mathcal{K}'^i$ of $\mathcal{K}^i$ having a subdivision Ψ^i of Φ^i as a subcomplex and such that the $\mathcal{K}'^i$-neighborhood of Ψ^i is normal. By reference to *Topology*, p. 91, it will be seen that both conditions may be fulfilled. Moreover we may assume the $\mathcal{K}$'s taken initially mutually exclusive, so that the closed $\mathcal{K}'$-neighborhoods introduced are all mutually exclusive. To simplify matters we designate henceforth these closed neighborhoods themselves by $\mathcal{K}^i$. Besides being polyhedral these neighborhoods have the following property (*loc. cit.*): if $\mathcal{B}^i = F(\mathcal{K}^i)$ then through every point P of $\mathcal{K}^i - \mathcal{B}^i - \Psi^i$ there passes a unique (open) segment resting on $\mathcal{B}^i$ and Ψ^i and varying continuously with P. We call these segments the *projecting segments* on $\mathcal{K}^i$.

† See *Annals of Mathematics*, vol. 32 (1931), p. 527.

36. Let us assign to each cell E_q^i of Φ^i one of its vertices A^{iq}, and let Ψ'^i be the first derived of Ψ^i. A unique simplicial transformation θ of Ψ'^i into Φ^i is determined by specifying that the vertices of Ψ'^i on E_q^i are all to be transformed by θ into A^{iq}. We shall designate by *projection* of Ψ'^i onto Φ^{i-1}, $\Phi^{i-2}, \cdots$, the simplicial transformation $\tau_{i-1}\theta, \tau_{i-2}\tau_{i-1}\theta, \cdots$. It is to be observed that θ need not be a fixed simplicial transformation, but merely any simplicial transformation of its type.

Let us specify for each i a definite first derived $\mathcal{K}'^i$. It determines a Ψ^i, and also a dual $\mathcal{K}^{*i}$ of $\mathcal{K}^i$. If Γ_{n+1} represents the fundamental cycle γ_n of M, then we have a definite n-cycle $\Gamma_n^i = \Gamma_{n+1} \cdot \Psi^i$ for each i. It is the subdivision induced by Ψ^i on the trace $\Gamma_{n+1} \cdot \Phi^i$ of Γ_{n+1}.

Now referring to *Topology*, Ch. IV, if $C^*{}_q^i$ is any subchain of $\mathcal{K}^{*i}$ there is a uniquely defined intersection-chain $C_s^i = \Gamma_n^i \cdot C^*{}_q^i$, $s = q + n - r$, and we have (*Topology*, p. 169, formula 18):

$$\Gamma_n^i \cdot C^*{}_q^i \rightarrow \Gamma_n^i \cdot F(C^*{}_q^i). \tag{36.1}$$

The intersection and its boundary are both subchains of Ψ'^i and hence they have a unique projection on any Φ^j, $j \leqq i$. Moreover, since a projection is a simplicial transformation, the boundary of the projection is the projection of the boundary.

37. In the argument to follow, in addition to the customary associated chains c_q, C_{q+1}, it will be convenient to make a clear distinction between intersections of chains and traces. We shall therefore designate the former as usual by the "dot" product, and the trace of C_{q+1} of K on Φ^i by $\mathscr{C}_q^i$.

LEMMA. *Let $C^\alpha{}_{p+1}$, $C^\alpha{}_{n-p+1}$ $(\alpha = 1, 2, \cdots, s)$ be projection-chains of K representing chains c_p^α, $c^\alpha{}_{n-p}$ which do not intersect one another's boundaries so that $(c_p^\alpha \cdot c^\alpha{}_{n-p})$ is well defined. Suppose that the traces $\mathscr{C}_p^{\alpha i}$ have the following property: whatever h there is a $k_\alpha > h$ such that $\mathscr{C}_p^{\alpha i}$ is the projection of an intersection-chain $\Gamma_n^{k_\alpha} \cdot C^{*\alpha k_\alpha}_{r-n+p}$ where*

$$\sum_\alpha \left(C^{*\alpha k_\alpha}_{r-n+p} \cdot \mathscr{C}^{\alpha k_\alpha}_{n-p}\right) = \lambda \tag{37.1}$$

is independent of h. Then,

$$\sum_\alpha \left(c_p^\alpha \cdot c^\alpha{}_{n-p}\right) = \lambda. \tag{37.2}$$

Let Λ_p designate the Lemma as stated. We shall reduce Λ_p to Λ_{p-1}, hence to Λ_0, then prove Λ_0.

Dropping α for the present, designate $C^\alpha{}_{p+1}, \cdots$ by $C_{p+1} \cdots$, and let V^j, $c^j{}_{n-p+1}$ correspond as in No. 29 to c_{n-p}. We first decompose c_p into elements $\{c_q^a\}$ which are projection-chains whose diameters $< \epsilon$. We then ϵ-deform

c_{n-p} and the chains c^j_{n-p+1} simultaneously so as not to impair their relation to one another and to the V's, and also so that they meet only the elements $c_p{}^a$ of the decomposition of c_p. It is readily seen that all these conditions can be fulfilled with ϵ as small as we please. We choose it $< \frac{1}{2}\zeta$, where ζ is the same as in No. 29. As a consequence we now have a pair c_p, c_{n-p} whose intersection consists of a finite number of disjoined closed sets F^a, one on each $c_p{}^a$. Since $\operatorname{diam} F^a < \frac{1}{2}\zeta$, F^a will be covered by a certain set V, which we may call V^a, such that $d(F^a, M - V^a) > \frac{1}{2}\zeta$. Since the F's do not intersect we can find for each F^a an open set W^a such that $F^a \subset W^a \Subset V^a$, $\bar{W}^a \cdot \bar{W}^b = 0$ for $a \neq b$. Introduce the closed set $G = M - \Sigma W^a$ and let L be a fundamental projection-complex for G (No. 7) so that $G = \bar{L} \cdot M$. We now remove from C_{p+1} all the p-cells on the Φ's which are on L and also all their joining cells, and call C'_{p+1} the chain left, C''_{p+1} the chain removed and c'_p, c''_p the corresponding chains of M whose sum is c_p. We have

$$c_p = c'_p + c''_p, \quad F(c_p) \subset M - \Sigma V^a \subset M - \Sigma W^a \subset M - c'_p.$$

Hence $| F(c'_p) | \subset | c''_p |$, and by construction the two chains c'_p, c''_p have only boundary points in common. Therefore

$$(37.3) \qquad | c'_p | \cdot | c''_p | = F(c'_p).$$

On the other hand if $c'_p{}^a$ designates the part of c'_p on $\bar{W}^a$, we have

$$(37.4) \qquad c'_p = \Sigma c'_p{}^a, \quad | c'_p{}^a | \cdot | c'_p{}^b | = 0 \text{ for } a \neq b.$$

Therefore also

$$(37.5) \qquad | c'_p{}^a | \cdot | c''_p | = F(c'_p{}^a).$$

As a consequence of (37.4) (second relation), for i sufficiently large $| \mathscr{C}'_p{}^{ai} |$ and $| \mathscr{C}'^{bi} |$, $a \neq b$, will have Φ^i-neighborhoods without common cells, for otherwise we would have $d(| c'_p{}^a | \cdot | c'_p{}^b |) = 0$. We also know that by construction $\mathscr{C}'_p{}^{ai}$ and $\mathscr{C}''_p{}^i$ have no common p-cells. Combining with the construction of C'_{p+1}, C''_{p+1}, we have

$$(37.6) \qquad | F(C'^a_{p+1}) | \cdot | C''_{p+1} | = | F(C'^a_{p+1}) | \subset L.$$

38. Until further notice we shall impose upon the simplicial transformation θ of No. 36 the following additional restriction: whenever $E_q{}^i$ is a cell of $\mathscr{C}''_p{}^i$ with vertices on L, we choose one of these vertices as the A^{iq} for that cell, that is as the vertex of Φ^i into which θ is to transform all the vertices of Ψ'^i that are on $E_q{}^i$.

Consider now $\mathring{C}^{*k}_{r-n+p}$ and let C'^{*k}_{r-n+p} be the chain left on removing from it the cells which do not meet $\mathscr{C}'_p{}^k$. Due to the mode of separation of the

chains $\mathfrak{C}'^{ak}_p$, for k sufficiently high C'^{*k}_{r-n+p} will be a sum of disjoined chains C'^{*ak}_{r-n+p} consisting respectively of the cells which meet the chain $\mathfrak{C}'^{ak}_p$. Therefore C'^{*ak}_{r-n+p} meets $\mathfrak{C}'^{ak}_p$ but not $\mathfrak{C}'^{bk}_p$ for $b \neq a$.

Now observe that as regards the cells of $\Gamma_n{}^k \cdot C^{*k}_{r-n+p}$ that are on the chains $\mathfrak{C}'$ the transformation θ preserves the same properties as in the Lemma. However it now takes the p-cells on a $\mathfrak{C}''$ and transforms them like cells on an $F(\mathfrak{C}')$, i. e. into cells of dimension $< p$ so that the projections of their boundaries do not affect the projections of the chains $F(\Gamma_n{}^k \cdot C'^{*ak}_{r-n+p})$. It follows that as regards the effect on the intersections $\Gamma_n{}^k \cdot C'^{*ak}_{r-n+p}$ its performance is as before and that this chain is now projected into $\mathfrak{C}'^{ai}_p$. It follows also that $F(\Gamma_n{}^k \cdot C'^{*ak}_{r-n+p})$ is projected at the same time into $F(\mathfrak{C}'^{ai}_p)$. All this holds of course for k large enough, which is all that we need.

It follows from what precedes that we may replace the initial chain c_p by a set of chains c'^a_p whose boundaries behave in a manner similar to that imposed upon c_p by the Lemma.

39. Since $c'^a_p \subset V^a$ we have from our discussion of the index

$$(39.1) \qquad (c_p \cdot c_{n-p}) = \Sigma\, (c'^a_p \cdot F(c^a_{n-p+1})).$$

Similarly since $r - n$ is even by *Topology*, p. 169, formula (20),

$$(39.2) \qquad \begin{aligned}(C^{*k}_{r-n+p} \cdot \mathfrak{C}^k_{n-p}) &= \Sigma\, (C'^{*ak}_{r-n+p} \cdot \mathfrak{C}^{ak}_{d-u}) \\ &= (-1)^p \Sigma\, (F(C'^{*ak}_p) \cdot \mathfrak{C}^{ak}_{n-p+1}).\end{aligned}$$

Comparing these relations and bringing back the index α, we find

$$(39.3) \qquad \Sigma\, (F(C'^{a_\alpha k_\alpha}_{r-n+p}) \cdot \mathfrak{C}^{a_\alpha k_\alpha}_{n-p+1}) = (-1)^p \cdot \lambda,$$

and the proof of the Lemma is reduced to showing that

$$(39.4) \qquad \Sigma\, (F(c'^{a_\alpha}_p) \cdot c^{a_\alpha}_{n-p+1}) = (-1)^p \cdot \lambda,$$

the relations between corresponding chains being as for Λ_{p-1}. That is to say we have reduced Λ_p to Λ_{p-1}, and hence to Λ_0.

40. We take up Λ_0, and we shall in fact prove the somewhat more stringent result that Λ_0 holds with all the numbers k_α equal, i. e. with a single chain $c_0{}^a$. We may go as far as No. 39 in the same manner as previously. Referring to No. 37 we have to prove that when the diameters of the sets $c'_0{}^a$ are small enough, if there is a k arbitrarily high such that $\mathfrak{C}_0{}^{ai}$ is the projection of $\Gamma_n{}^k \cdot C^{*ak}_{r-n}$, then

$$(40.1) \qquad \Sigma\, t_a(C^{*ak}_{r-n} \cdot \Gamma_n{}^k) = \Sigma\, t_a(c_0{}^a \cdot \gamma_n).$$

This will follow if we can show that

(40.2) $$(C^{*ak}_{r-n} \cdot \Gamma_n{}^k) = (c_0{}^a \cdot \gamma_n).$$

In the first place we have (No. 24)

(40.3) $$(c_0{}^a) = (c_0{}^a \cdot \gamma_n) = (\mathscr{C}_0{}^{ai})$$

for i large enough. Also since $\mathscr{C}_0{}^{ai}$ is the projection of $\Gamma_n{}^k \cdot C^{*ak}_{r-n}$, we have

(40.4) $$(\mathscr{C}_0{}^{ai}) = (\Gamma_n{}^k \cdot C^{*ak}_{r-n}) = (C^{*ak}_{r-n} \cdot \Gamma_n{}^k),$$

from which (40.2) and hence (40.1) follow. This proves Λ_0 and hence also the Lemma.

41. An important application of the Lemma is the proof, still lacking, of formula (34.7), and hence of property (23 f), for the index. For take first $p = 0$, and $C^a{}_{p+1} = C_1 =$ a chain made up of a single projecting line whose traces $\mathscr{C}_0{}^i$ are vertices A^i of the complexes Φ^i such that for i above a certain value A^i is the vertex of an E_n of Φ^i. Then taking $k_a = i$, $C^{*a}_{r-n} = E^*{}_{r-n}$ the cell of $\mathscr{K}^i$ dual to E_n, $\mathscr{C}^{ak_a}_{r-n} = \Gamma_n{}^i$ and orientations as in *Topology*, Ch. IV, the condition of the Lemma is fulfilled with a single $c_0 = x$ and a single $c_n = \gamma_n$. Therefore

(41.1) $$(x \cdot V_n) = (E^*{}_{r-n} \cdot E_n) = +1.$$

Choose now $p = n$ and $C^a{}_{p+1} = C_{n+1} = \Gamma_{n+1}$ the projection-chain defining γ_n, $k_a = i$, $C^{*ak_a}_{r-n+p} = C^*{}_r{}^i$, the sum of the cells of $\mathscr{K}^{*i}$ oriented like the spaces S_r, $\mathscr{C}^{ak_a}_{n-p} =$ the same vertex A^i as previously. This time the conditions of the Lemma are again fulfilled and we find

(41.2) $$(\gamma_n \cdot x) = (C^*{}_r{}^i \cdot A^i) = +1 = (x \cdot \gamma_n),$$

which is the result that we required.

42. From the Lemma to the duality formula for the Betti-numbers is but a step. Let $\gamma^a{}_{n-p}$, $\alpha = 1, 2, \cdots, R_{n-p}$, be a base for the $(n-p)$-cycles consisting of irreducible projection-cycles (Theorem II). Let $\Gamma^a{}_{n-p+1}$, $\mathscr{G}^{ah}_{n-p}$ be the representative cycle mod Φ, and trace on Φ^h, associated with $\gamma^a{}_{n-p}$. Then for h above a certain value the cycles $\mathscr{G}^{ah}_{n-p}$ are independent on Φ^h, hence also independent on $\mathscr{K}^h - \mathscr{B}^h$. For if say

$$\mathscr{K}^h - \mathscr{B}^h \supset C_{n-p+1} \to \Sigma\, t_a \mathscr{G}^{ah}_{n-p},$$

we could slide down C_{n-p+1} along the projecting segments on $\mathscr{K}^h - \mathscr{B}^h - \Phi^h$ onto Φ^h, and obtain a chain on Φ^h

$$C'_{n-p+1} \to \Sigma\, t_a \mathscr{G}^{ah}_{n-p}$$

so that the cycles $\mathscr{G}^{ah}_{n-p}$ would not be independent on Φ^h.

As a consequence of the independence of these cycles on $\mathcal{K}^h - \mathcal{B}^h$, $\mathcal{K}^h$ contains a cycle mod $\mathcal{B}^h$, $\mathcal{G}^{ah}_{r-n+p}$ whose cells intersecting Φ^h consist of cells of the dual $\mathcal{K}^{*h}$, and such that (*Topology*, pp. 140, 174).

$$(42.1) \qquad (\mathcal{G}^{ah}_{r-n+p} \cdot \mathcal{G}^{\beta h}_{n-p}) = \delta_{\alpha\beta}.$$

Consider now the projection $\Gamma_p{}^i$ of all cycles $t\Gamma_n{}^h \cdot \mathcal{G}^{ah}_{r-n+p}$ (α fixed) on a definite Φ^i. As far as the intersections with Φ^h go the chain $\mathcal{G}^{ah}_{r-n+p}$ is a $C^{*h}{}_{r-n+p}$. With $\Gamma_p{}^i$ we associate the numbers $(t\delta_{\alpha\beta})$ and if $\Gamma'_p{}^i$ corresponds to t' and the numbers $(t'\delta_{\alpha\beta})$, we associate with $s\Gamma_p{}^i + s'\Gamma'_p{}^i$ the numbers $((st + s't')\delta_{\alpha\beta})$. In this manner if $\mathfrak{M}^i$ is the modulus generated by the cycles $\Gamma_p{}^i$, there corresponds to each member of $\mathfrak{M}^i$ a definite set $(t\delta_{\alpha\beta})$. Clearly members corresponding to $t = 0$ give rise to a submodulus $\mathfrak{N}^i$ of $\mathfrak{M}^i$. Also by construction the moduli $\mathfrak{M}^i, \mathfrak{N}^i$ are in the very relationship demanded by Theorem I. Therefore there exists a projection-sequence $\{\Gamma_p{}^{ai}\}$ such that the cycle $\Gamma_p{}^{ai}$ is a member of $\mathfrak{M}^i$ corresponding to $t = 1$. This sequence gives rise to a projection-cycle mod Φ, $\Gamma^a{}_{p+1}$, which defines a normal cycle $\gamma_p{}^a$. Owing to (42.1) and to the mode of defining the moduli $\mathfrak{M}^i$, we have by the Lemma

$$(42.2) \qquad (\gamma_p{}^a \cdot \gamma^\beta{}_{n-p}) = \delta_{\alpha\beta}.$$

Hence (No. 23 property e) the cycles $\gamma_p{}^a$ are independent and therefore $R_p \geqq R_{n-p}$. Similarly $R_p \leqq R_{n-p}$ and therefore we have proved Poincaré's duality relation for an absolute n-manifold:

$$(42.3) \qquad R_p(M) = R_{n-p}(M).$$

43. *Extension to open manifolds.* Take first an open M_n and let U be an open subset of M whose closure $\bar{U}$ is self-compact. Then if $V \Subset U$, $\bar{V}$ is likewise self-compact. As the manifold conditions hold over U we may apply Theorem VII with the following slight restrictions: $c_p \subset V, \epsilon < d(V, M - U)$. From this we conclude, as in No. 28, that there are at most finite numbers: (a) of absolute p-cycles of U independent mod $M - V$; (b) of p-cycles of U mod $M - U$, independent mod $M - V$. We can then show as in the preceding number that the two numbers are equal.

The sequences of open sets $\{U^i\}$ such that $U^{i+1} \Subset U^i$, $\Pi U^i = 0$, may serve to define the different types of ideal elements as we have done in *Topology*, Ch. VII. In the terminology there used let Λ be the total ideal element, and let $\mathcal{L}^1, \mathcal{L}^2$ designate complementary closed and open ideal elements. Let also L be any closed subset of M which $\supset \Lambda$ and let L^1 be any closed subset of L with $\mathcal{L}^1$ for ideal element. Then if $L^2 = L - L^1$, $\mathcal{L}^2$ will be the ideal element of L^2. By means of properties (a), (b), (c), and by unimportant adaptations of the treatment in *Topology*, Ch. VII, § 3, we prove:

FUNDAMENTAL DUALITY THEOREM. *Let Γ_p, G_{n-p} be associated cycles of the dual types $M - L^1 \bmod L^2$ and $M - L^2 \bmod L^1$ in any ring of rational coefficients forming a field. There exists two associated dual bases $\{\Gamma_p{}^{\alpha}\}$, $\{G^{\beta}{}_{n-p}\}$ made up of true normal cycles whose indices satisfy the relations*

$$(\Gamma_p{}^{\alpha} \cdot G^{\beta}{}_{n-p}) = \delta_{\alpha\beta}. \tag{43.1}$$

Whatever Γ_p, G_{n-p} we have

$$\Gamma_p = \Sigma\,(\Gamma_p \cdot G^{\alpha}{}_{n-p}) \cdot \Gamma_p{}^{\alpha}, \tag{43.2}$$

$$G_{n-p} = \Sigma\,(\Gamma_p{}^{\alpha} \cdot G_{n-p}) \cdot G^{\alpha}{}_{n-p}, \tag{43.3}$$

and the Betti-numbers satisfy the duality relations

$$R_p(M - L^1, L^2) = R_{n-p}(M - L^2, L^1). \tag{43.4}$$

In particular: (a) *when $L^1 = \Lambda$, $L^2 = 0$* we have

$$R_p(M - \Lambda) = R_{n-p}(M, \Lambda), \tag{43.5}$$

where the Betti-numbers refer at the left to the finite cycles and at the right to the infinite cycles; (b) *when $\Lambda = 0$, i. e. when M is absolute, the bases are finite and the duality relation reduces to that of Poincaré*

$$R_p(M) = R_{n-p}(M). \tag{43.6}$$

These results hold also when M consists of a countable aggregate of circuits.

The last part of the statement, regarding an M consisting of a countable aggregate of circuits, is an immediate consequence of the following: when $M = \Sigma\, M^i$, where M^i is a connected n-manifold and $M^i \cdot M^j = 0$, then the p-th homology group of M is the direct sum of those of the manifolds M^i (the groups are assumed written additively).

SYSTEMS OF ALGEBRAIC DIFFERENCE EQUATIONS.

By J. F. Ritt and J. L. Doob.*

The object of this paper is to derive an analogue, for systems of algebraic difference equations, of the fundamental theorem in the theory of systems of algebraic differential equations developed by one of us.† We introduce the notion of irreducible system of algebraic difference equations, and show that every system of such equations is equivalent to a finite set of irreducible systems. It will possibly strike one as curious that so general a result should be obtainable at a time when existence theorems for non-linear difference equations are almost entirely lacking.

Although our proof resembles greatly that for differential equations, there are also essential differences. These arise out of the circumstance that the derivative of a polynomial in several functions involves the derivatives of the functions linearly, while no corresponding result holds for the operation of differencing.

Fields, Forms, Solutions.

1. Let $\mathfrak{A}$ be an open region, in the plane of the complex variable x, which contains $x+1$ when it contains x. A set $\mathfrak{F}$ of functions meromorphic in $\mathfrak{A}$ and not all zero, will be called a field when both of the following conditions are satisfied:

(a) If $f(x)$ is in $\mathfrak{F}$, $f(x+1)$ is in $\mathfrak{F}$.

(b) If $f(x)$ and $g(x)$ are in $\mathfrak{F}$,

$$f \pm g; \quad fg; \quad f/g, \quad (g \neq 0)$$

are in $\mathfrak{F}$.

By a *form*, we mean a polynomial in a finite number of the symbols $y_i(x+j)$, $i=1,\cdots,n$; $j=0,1,2,\cdots$, with coefficients which are functions of x meromorphic in $\mathfrak{A}$. The integer n will be considered fixed throughout the discussion.

All forms appearing in this paper will be understood to have coefficients belonging to a given field $\mathfrak{F}$.

Throughout our work, capital italic letters will denote forms.

---

* National Research Fellow.

† Ritt, "Differential equations from the algebraic standpoint," *Colloquium Publications of the American Mathematical Society*, vol. XIV (1932). The present paper is complete in itself.

Let F be any form. Let $\mathcal{L}$ be a polygonal line without double points, lying in $\mathfrak{A}$, and such that $x+1$ is on $\mathcal{L}$ whenever x is. Let $y_1(x), \cdots, y_n(x)$ be a set of functions, analytic on $\mathcal{L}$, which cause F to vanish. The entity composed of $\mathcal{L}$ and of $y_1(x), \cdots, y_n(x)$ will be called a *solution* of F (or of $F=0$). Of course, $\mathcal{L}$ may be different for different solutions of F.*

Let Σ be any finite or infinite system of forms. By a *solution* of Σ we mean a common solution of the forms of Σ. The totality of solutions of Σ will be called the *manifold* of Σ. If Σ_1, and Σ_2 are systems such that every solution of Σ_1 is a solution of Σ_2, we shall say that Σ_2 *holds* Σ_1.†

Rank of Forms.

2. By a *transform* of the function $y_i(x)$, we shall mean any function $y_i(x+r)$ with r a non-negative integer. We shall call r the *order* of the transform.

By the *class* of a form F, if F actually involves the unknowns, we shall mean the greatest value of j such that some transform of y_j is present in F. If F is merely a function of x, F will be said to be of class zero.

If F is of class $p>0$, we shall understand by the *order* of F the order of the highest transform of y_p which appears in F.

Let F_1 and F_2 be two forms. If F_2 is of higher class than F_1, we shall say that F_2 is *of higher rank* or *higher* than F_1.

Let F_1 and F_2 be of the same class $p>0$. We shall say that F_2 is higher than F_1 if either

(a) F_2 is of higher order than F_1,

or

(b) F_1 and F_2 are of the same order, say q, and F_2 is of higher degree than F_1 in $y_p(x+q)$.

Two forms for which no difference in rank is created by what precedes will be said to be of the same rank. For instance, all forms of class 0 are of the same rank.

If F_1 is higher than F_2 and F_2 higher than F_3, then F_1 is higher than F_3.

We prove the following lemma:

Lemma. *In every aggregate of forms, there is a form which is not higher than any other form of the aggregate.*

If there are forms of class zero in the aggregate, any such form will serve

* This definition of solution can be broadened considerably. With suitable explanation, $\mathcal{L}$ may be allowed to intersect itself.

† If Σ_1 has no solutions, every system will be said to hold Σ_1.

as one of lowest rank. If not, let $p > 0$ be the least of the classes of the forms. From among all forms of class p in the aggregate we select those whose order is a minimum, say q. From the forms just selected we take one which is of minimum degree in $y_p(x+q)$. This form fulfills our conditions.

Ascending Sets.

3. Let A_1 be a form of class $p > 0$ and of order q. A form A_2 will be said to be *reduced with respect to* A_1 if $A_2 = 0$ or if $A_2 \neq 0$ and the degree of A_2 in every $y_p(x+r)$ with $r \geq q$ is less than the degree of A_1 in $y_p(x+q)$.*

The system

$$A_1, A_2, \cdots, A_r \tag{1}$$

will be called an *ascending set* if either

(a) $r = 1$ *and* $A_1 \neq 0$

or

(b) $r > 1$, A_1 *is of class higher than* 0, *and, for* $j > i$, A_j *is of higher rank than* A_i *and reduced with respect to* A_i.†

The ascending set (1) will be said to be of *higher rank* than the ascending set

$$B_1, B_2, \cdots, B_s$$

if either

(a) *There is a* j, *exceeding neither* r *nor* s, *such that* A_i *and* B_i *are of the same rank for* $i < j$ *and that* A_j *is higher than* B_j.‡

or

(b) $s > r$ and A_i *and* B_i *are of the same rank for* $i \leq r$.

Two ascending sets for which no difference in rank is created by what precedes will be said to be of the same rank. For such sets, $r = s$ and A_i and B_i are of the same rank for every i. The above ordering of ascending sets is easily seen to be transitive, but this fact will not be used in what follows.

We prove the following lemma:

Lemma. *Every finite or infinite aggregate of ascending sets contains an ascending set whose rank is not higher than that of any other ascending set in the aggregate.*

---

* This definition is materially different from the corresponding definition for differential forms.

† Note that A_{i+1} need not be of higher class than A_i. This is a respect in which our ascending sets differ from ascending sets of differential forms.

‡ If $j = 1$, this is to mean that A_1 is higher than B_1.

Among the ascending sets in the aggregate, there are, by the Lemma of § 2, certain ones whose first forms are of a least rank. Let σ_1 be the totality of such ascending sets. If the sets in σ_1 all consist of one form, any set in σ_1 will serve as the set whose existence was to be proved. Suppose that σ_1 contains sets which have more than one form. From among all such sets in σ_1 we select those whose second forms have a least rank, and denote the totality of the sets selected by σ_2. We continue in this fashion. If we meet a σ_m whose sets have exactly m forms, any set of σ_m will satisfy the requirements of the lemma. We shall show that such a σ_m eventually presents itself. Taking any σ_m which is met, let the m-th form in any of its sets be of class p, order q and of degree t in $y_p(x+q)$. Suppose now that σ_{m+1} exists and let the $(m+1)$-th forms in its sets be of class p', order q' and degree t' in $y_{p'}(x+q')$. Suppose that $p'=p$. Having regard to what it means for one form to be of higher rank than, and reduced with respect to, a second, we see that $q'>q$ and $t'<t$. On this basis, if σ_{m+t} exists, the class of the $(m+t)$-th forms in its sets must exceed p. As all forms are of class not exceeding n, the process of forming the systems σ_m must terminate after a finite number of steps. This proves the lemma.

Basic Sets.

4. Let Σ be any finite or infinite system of forms, not all zero. There exist ascending sets in Σ; for instance, every non-zero form of Σ is an ascending set. Among all ascending sets in Σ, there are, by § 3, certain ones which have a least rank. Any such ascending set in Σ will be called a *basic set* of Σ.

If A_1, in (1), is of class greater than zero, a form F will be said to be *reduced with respect to the ascending set* (1) if F is reduced with respect to A_i, $i=1,\cdots,r$.

Let Σ be a system for which (1), with A_1 of class greater than 0, is an ascending set. Let F be a non-zero form which is reduced with respect to (1). Let $\Sigma+F$ denote the system composed of F and of the forms of Σ. We shall prove that *the basic sets of $\Sigma+F$ are of lower rank than those of Σ.*

It will suffice to show that $\Sigma+F$ contains an ascending set lower than (1). If F is lower than A_1, F is an ascending set lower than (1). If not, since F is reduced with respect to A_1, F must be of higher rank than A_1. Then, if F is lower than A_2, the ascending set A_1, F is lower than (1). We continue, terminating with the possibility that F is higher than A_r, in which case $A_1,\cdots,A_r, F$ is an ascending set in $\Sigma+F$ lower than (1).

In the same way we can show that *if* (1) *with A_1 of class higher than zero is a basic set of a system Σ, then Σ cannot contain a non-zero form reduced with respect to* (1).

REDUCTION.

5. If A is any form, the form obtained on replacing x by $x+m$, where m is a non-negative integer, in the coefficients of A and in the $y_i(x+j)$ appearing in A, will be called the *m-th transform of A.*

If A is of class $p>0$ and of order q, the coefficient of the highest power of $y_p(x+q)$ in A will be called the *initial* of A. The initial of A is lower than A.

We consider an ascending set

$$A_1, \cdots, A_r \tag{2}$$

with A_1 of class higher than zero. We prove the following lemma:

LEMMA. *Let G be any form. There exists a form J which is a product of powers of the initials of the A_i in* (2) *and of transforms of those initials, such that, when a suitable linear combination of the A_i and of a certain number of their transforms, with forms for coefficients, is subtracted from JG, the remainder, R, is reduced with respect to* (2).

Let I_i denote the initial of A_i, $i=1,\cdots,r$.

We represent the m-th transform of any form A by $A(x+m)$.

We may limit ourselves to the case in which G is not reduced with respect to (2). Let j be the greatest value of i such that G is not reduced with respect to A_i. Let A_i be of class p and order q. Let $y_p(x+h)$ be the highest transform of y_p appearing in G. Then $h \geqq q$. Let us suppose that $h>q$. If $k=h-q$, then $A_j(x+k)$ will be of order h, with $I_j(x+k)$ for initial. Using the algorithm of division, we determine a non-negative integer v_1 such that

$$[I_j(x+k)]^{v_1} G = C_1 A_j(x+k) + D_1 \tag{3}$$

where D_1 is either zero or has a lower degree in $y_p(x+h)$ than A_j has in $y_p(x+q)$. For uniqueness of procedure we take v_1 as small as possible.

Let z be any form of the type $y_i(x+s)$ which is higher than $y_p(x+h)$. We shall show that D_1, if not zero, is not of higher degree than G in z. Let this be untrue. Then, as $I_j(x+k)$ and $A_j(x+k)$ do not involve z, C_1 must involve z in the same power in which D_1 does. Then $C_1A_j(x+k)$ contains terms involving z and $y_p(x+h)$ which can be balanced neither by D_1 nor by the first member of (3). This proves our statement.

Thus if $j<r$, D_1 is reduced with respect to $A_{j+1},\cdots,A_r$.

If D_1 is not reduced with respect to A_j, we find a relation

$$[I_j(x+k-1)]^{v_2} D_1 = C_2 A_j(x+k-1) + D_2$$

where D_2 is zero or has a lower degree in $y_p(x+h-1)$ than A_j has in $y_p(x+q)$. If $D_2 \neq 0$, its degree in any $y_i(x+s)$ higher than $y_p(x+h-1)$ does not exceed that of D_1. For uniqueness we take v_2 as small as possible.

Continuing, we find a D_u which is reduced with respect to $A_j, \cdots, A_r$. Evidently D_u differs from some J_1G, where J_1 is a product of powers of I_j and its transforms, by a linear combination of A_j and its transforms.

If D_u is not reduced with respect to (2), we give it the treatment accorded to G. For some $l < j$, there is a form J_2 which is a product of powers of I_l and its transforms, such that J_2D_u differs by a linear combination of A_l and its transforms from a form D_v which is reduced with respect to $A_l, \cdots, A_r$. Evidently, J_1J_2G exceeds D_v by a linear combination of A_j, A_l and their transforms.

Continuing, we reach a form R as described in the statement of the lemma. Our procedure determines a unique R. We call this R the *remainder of G with respect to the ascending set* (2).

Completeness of Infinite Systems.

6. In §§ 6-8, we prove the following lemma:

Lemma. *Every infinite system of forms in $y_1, \cdots, y_n$ has a finite subsystem whose manifold is identical with that of the infinite system.*

An infinite system whose manifold is identical with that of one of its finite subsystems will be called *complete.** Infinite systems which are not complete will be called *incomplete.* In what follows we assume the existence of incomplete systems and force a contradiction.

7. We prove the following lemma:

Lemma. *Let Σ be an incomplete system. Let $F_1, \cdots, F_s$ be such that, by multiplying each form in Σ by some product of powers of the F_i and their transforms, a system Λ is obtained which is complete.† Then at least one of the systems $\Sigma + F_i$, $i = 1, \cdots, s$, is incomplete.*

Suppose that every system $\Sigma + F_i$ is complete. Then, for every i, there is a finite subsystem Φ_i of $\Sigma + F_i$ which has the same manifold as the latter system. As Φ_i may evidently be replaced by any finite subsystem of $\Sigma + F_i$ which contains Φ_i, we may suppose Φ_i, for every i, to be of the type

---

* If some finite subsystem has no solutions, the system will be considered complete.

† The product of powers may, of course, be different for different forms of Σ.

$$(4) \qquad F_i, A_1, \cdots, A_q$$

with the set

$$(5) \qquad A_1, \cdots, A_q$$

independent of i. We may, furthermore, enlarging (5) sufficiently, assume that the forms of Λ obtained from (5) by the above described multiplications form a system with the same manifold as Λ.

Let L, in Σ, not hold (5). (§ 1). Now the product of L by some product of powers of the F_i and their transforms is in Λ, and holds (5). A form has the same manifold as any of its transforms. Hence

$$F_1 \cdot F_2 \cdots F_s \cdot L$$

holds (5). This means that certain solutions of (5) which are solutions of $F_1 \cdots F_s$ are not solutions of L. Then some F_i has a solution in common with (5) which is not a solution of L. In other words, there is an i for which L does not hold (4). This proves the lemma.

8. Let us consider the totality of incomplete systems of forms in $y_1, \cdots, y_n$. According to § 3, there is one of them, Σ, whose basic sets (§ 4), are not higher than those of any other incomplete system. Let (2) be a basic set of Σ. Then A_1 must be of class greater than zero, else A_1 would have no solutions and Σ would be complete.

For every form of Σ not in (2), let a remainder with respect to (2) be found as in § 5. Let Λ be the system composed of the forms of (2) and of the products of the forms of Σ not in (2) by the power products of the I_i and their transforms used in forming the remainders. Let Ω be the system composed of (2) and of the remainders of the forms of Σ not in (2).

Now Ω must be complete. If not, it would certainly have non-zero forms not in (2). Since such forms would be reduced with respect to (2), then (2) could not be a basic set of Ω (§ 4). Then the basic sets of Ω would be lower than (2) and Σ would not be an incomplete system with lowest basic sets.

If H is a form of Λ not in (2) and R the corresponding form in Ω, then H and R have the same solutions in common with (2). This means that Λ and Ω have the same manifold and that Λ is complete.

The lemma of § 7 shows us now that some $\Sigma + I_i$ is incomplete. But, for every i, I_i is distinct from zero and reduced with respect to (2). Then, by § 4, the basic sets of every $\Sigma + I_i$ are of lower rank than (2). This proves the fundamental lemma stated in § 6.

Irreducible Systems.

9. A system Σ will be said to be *reducible* if there exist two forms, G and H, such that neither G nor H holds Σ while GH holds Σ. Systems which are not reducible will be called *irreducible*.

For instance, let Σ, in the single unknown y, be the first member of the difference equation

$$[y(x+1)-y(x)]^2-[y(x+1)+y(x)]=0. \tag{6}$$

Let $\mathfrak{F}$ be the field of all constants. We observe that the first member of (6), considered as a polynomial in $y(x)$ and $y(x+1)$, is algebraically irreducible in $\mathfrak{F}$.

We shall solve (6). Let x be replaced by $x+1$ in (6) and let (6) be subtracted from the resulting equation. We find

$$[y(x+2)-y(x)]\,[y(x+2)-2y(x+1)+y(x)-1]=0. \tag{7}$$

If

$$y(x+2)-2y(x+1)+y(x)=1 \tag{8}$$

then, since the first member of (8) is the first difference of $y(x+1)-y(x)$, we must have

$$y(x+1)-y(x)=x+\phi(x), \tag{9}$$

with $\phi(x)$ periodic and of period unity. From (6) and (9), we find

$$y(x+1)+y(x)=[x+\phi(x)]^2, \tag{10}$$

and (9) and (10) give

$$y(x)=\frac{[x+\phi(x)]^2-[x+\phi(x)]}{2}. \tag{11}$$

If

$$y(x+2)-y(x)=0, \tag{12}$$

$y(x)$ is periodic and of period 2. Now

$$y(x)=\frac{y(x)+y(x+1)}{2}+\frac{y(x)-y(x+1)}{2}. \tag{13}$$

The first fraction in the second member of (13) is of period unity, whereas the second fraction is multiplied by -1 when x is increased by unity. We may then write

$$y(x)=\phi_1(x)+e^{\pi ix}\phi(x)$$

with ϕ_1 and ϕ of period unity. Again, (6) gives

$$4e^{2\pi ix}[\phi(x)]^2=2\phi_1(x)$$

so that

$$(14) \qquad y(x) = 2e^{2\pi ix}[\phi(x)]^2 + e^{\pi ix}\phi(x).$$

Thus, the solutions of (6) are given by (11) and (14).

Now the solution $y = (x^2 - x)/2$, belonging to (11), does not annul the first factor in (7), while the solution $y = 0$, belonging to (14), does not annul the second factor.

Thus Σ is reducible. Let Σ_1 and Σ_2 be obtained by adjoining to Σ the first and second factors in (7) respectively. Obviously the manifold of Σ consists of the combined manifolds of Σ_1 and Σ_2.

We shall prove that Σ_1 is irreducible.

Let GH hold Σ_1. Let G_1 and H_1 be respectively the remainders of G and H with respect to

$$(15) \qquad y(x+2) - y(x).$$

As the initial of (15) is unity, G and G_1 have the same solutions in common with Σ_1; similarly for H and H_1. Also G_1 and H_1 are at most of order unity. If we can show that one of G_1, H_1 is divisible by the form in Σ, call it A, we shall know that one of G, H holds Σ_1. Suppose that G_1 and H_1 are not divisible by A. As A is algebraically irreducible, the resultant R of A and G_1H_1 with respect to $y(x+1)$, is not zero. Also R, like G_1H_1, holds Σ_1. But R is a polynomial in $y(x)$ alone and thus cannot admit the totality of solutions (14), which depends on an arbitrary function. This proves that Σ_1 is irreducible. Similarly Σ_2 is irreducible.

The Decomposition Theorem.

10. A system Σ will be said to be *equivalent to the set of systems* $\Sigma_1, \cdots, \Sigma_s$ if Σ holds every Σ_i, while every solution of Σ is a solution of some Σ_i. Thus, two systems with the same manifold are equivalent to each other.

We prove the following theorem:

Theorem. *Every system of forms is equivalent to a finite set of irreducible systems.*

Let the theorem be false for some system Σ. Then Σ is reducible. Let G_1 and H_1 be such that G_1H_1, but neither G_1 nor H_1, holds Σ. Then Σ is equivalent to the set

$$(16) \qquad \Sigma + G_1, \qquad \Sigma + H_1.$$

Evidently at least one of the systems in (16) has the property that it is not equivalent to a finite set of irreducible systems. Let $\Sigma + G_1$ have this property. We then find a G_2, which does not hold $\Sigma + G_1$, such that $\Sigma + G_1$

$+ G_2$ has the same property. Continuing, we find a G_p for every p. Then the system Ψ composed of

$$\Sigma,\ G_1,\ G_2, \cdots, G_p, \cdots$$

is incomplete. For if Ψ held

$$\Phi + G_{i_1} + \cdots + G_{i_q}$$

with Φ a finite subsystem of Σ and $i_1 < \cdots < i_q$, then Ψ would hold

$$\Sigma + G_1 + \cdots + G_{i_q}. \tag{17}$$

This cannot be, since G_{i_q+1} does not hold (17). This proves our theorem.

Uniqueness of Decomposition.

11. Let a system Σ be equivalent to the set of irreducible systems

$$\Sigma_1, \cdots, \Sigma_s. \tag{18}$$

We shall suppose, suppressing certain of the Σ_i, if necessary, that no Σ_i holds a Σ_j with $j \neq i$. We can then prove that *the decomposition* (18) *is essentially unique.* That is, *if $\Omega_1, \cdots, \Omega_t$ is a second decomposition of Σ into irreducible systems, none of which holds any other, then $t = s$ and every Ω_i is equivalent to some Σ_j.*

We shall show that there is some Ω_i which holds Σ_1. If there were not, then each Ω_i would have a form which would not hold Σ_1. Such forms being selected, their product would hold each Ω_i, consequently Σ, thus Σ_1. This is impossible if Σ_1 is irreducible and none of the forms holds Σ_1.

Then let Ω_1 hold Σ_1. Now Ω_1, similarly, must be held by some Σ_i, which must be Σ_1, since no Σ_i with $i \neq 1$ holds Σ_1. Thus Σ_1 and Ω_1 are equivalent. The uniqueness is proved.

Columbia University,
New York, N. Y.

THE CONVERGENCE OF SOME NON-LINEAR PROCESSES OF APPROXIMATION.*

By Dunham Jackson.

1. **Introduction.** Let $f(x)$ be a positive continuous function of period 2π, and let an approximation be sought for it in the form $e^{T_n(x)}$, where $T_n(x)$ is a trigonometric sum of the n-th order. If $T_n(x)$ is any sum approximating $\log f(x)$, naturally $e^{T_n(x)}$ gives some sort of approximation to $f(x)$. This paper is concerned specifically with the properties of sums $T_n(x)$ chosen so as to minimize the integral

$$(1) \qquad \int_{-\pi}^{\pi} \rho(x) \mid f(x) - e^{T_n(x)} \mid^m dx,$$

where $\rho(x)$ is a given non-negative summable weight function and m a given positive exponent. A significant feature of the problem is the fact that the approximating function depends non-linearly on the fundamental functions $\cos kx$ and $\sin kx$ in terms of which it is expressed, and the particular form chosen, in spite of its simplicity, introduces complications which were not encountered in an earlier paper by the writer with a similar title.† It will be shown nevertheless that under appropriate hypotheses the minimum problem has a solution (which to be sure is not shown to be unique), and that with more restrictive hypotheses the approximating functions $e^{T_n(x)}$ converge uniformly toward $f(x)$ as n becomes infinite. (Under the conditions imposed convergence of $e^{T_n(x)}$ toward $f(x)$ and convergence of $T_n(x)$ toward $\log f(x)$ are equivalent; the point is that the criterion defining $T_n(x)$ in the first place is altogether different from that which would be set up by requiring that an integral in terms of $|\log f(x)) - T_n(x)|$ be a minimum.)

A concluding section will give a brief discussion of the corresponding problem of polynomial approximation.

2. **Theorems of existence and convergence for function having a continuous derivative.** It will be assumed throughout this section that $m \geqq 1$, and that the function $f(x)$, continuous, of period 2π, and everywhere positive, has a continuous derivative for all values of x. The hypotheses imply of course that $f(x)$ has a positive minimum; let $h > 0$ be its minimum, M its

* Presented to the American Mathematical Society December 29, 1932.

† "Some non-linear problems in approximation," *Transactions of the American Mathematical Society*, Vol. 30 (1928), pp. 621-629.

maximum, and λ an upper bound for the absolute value of its derivative. Since the function

$$\phi(x) \equiv \log f(x)$$

has everywhere a continuous derivative, there exist * trigonometric sums $t_n(x)$ such that $\lim_{n\to\infty} n\epsilon_n = 0$, if ϵ_n is for each n the maximum of $|\phi(x) - t_n(x)|$. By the mean value theorem,

$$f(x) - e^{t_n(x)} = e^{\phi(x)} - e^{t_n(x)} = [\phi(x) - t_n(x)]e^{\xi_n(x)},$$

where $\xi_n(x)$ is intermediate in value between $\phi(x)$ and $t_n(x)$. Since the sums $t_n(x)$ uniformly approach $\phi(x)$ they are uniformly bounded, the functions $\xi_n(x)$ and $e^{\xi_n(x)}$ are uniformly bounded, and $|f(x) - e^{t_n(x)}|$ does not exceed a constant multiple of ϵ_n.

The functions $f(x)$ and $\rho(x)$ and the exponent $m \geqq 1$ being given, let γ_n be the greatest lower bound of the integral (1) as $T_n(x)$ ranges over all trigonometric sums of the n-th order; it is not assumed as yet that this lower bound is a minimum actually attained. It is clear that $\gamma_n \leqq k\epsilon_n{}^m$, where ϵ_n has the meaning given to it in the preceding paragraph and k is independent of n, and hence that

$$\lim_{n\to\infty} n^m\gamma_n = 0. \tag{2}$$

Let it be assumed throughout the rest of this section that the weight function $\rho(x)$ has a positive lower bound: $\rho(x) \geqq v > 0$ for all values of x.

Let $G > 0$ be the larger of the numbers $|\log h|$, $|\log M|$, the trivial case $f(x) \equiv 1$ being ruled out. For an arbitrary $T_n(x)$, let g denote the value of the integral (1). It will be shown that for n sufficiently large all sums $T_n(x)$ of the n-th order for which $g \leqq 2\gamma_n$ are subject to the inequality $|T_n(x)| < 4G$, for all values of x.

Let μ be the maximum of $|T_n(x)|$, and let it be supposed that $\mu \geqq 4G$. Let x_0 be a value of x such that $|T_n(x_0)| = \mu$. By Bernstein's theorem, $|T'_n(x)| \leqq n\mu$ everywhere, and

$$|T_n(x) - T_n(x_0)| \leqq n\mu|x - x_0|.$$

For $|x - x_0| \leqq 1/(2n)$,

$$|T_n(x) - T_n(x_0)| \leqq \tfrac{1}{2}\mu, \qquad |T_n(x)| \geqq \tfrac{1}{2}\mu \geqq 2G.$$

Throughout this interval, then, $e^{T_n(x)} \geqq e^{2G}$, or else $e^{T_n(x)} \leqq e^{-2G}$, one or the

* See e. g. D. Jackson, "The theory of approximation," *American Mathematical Society Colloquium Publications*, New York, Vol. 11 (1930), p. 12, Theorem IV, Corollary.

other of the indicated relations holding throughout the whole interval, according as $T_n(x_0) = \mu$ or $-\mu$. On the other hand,

$$e^{-G} \leqq h \leqq f(x) \leqq M \leqq e^{G}$$

everywhere. Consequently, for $|x - x_0| \leqq 1/(2n)$, and so throughout an interval of length $1/n$,

$$|f(x) - e^{T_n(x)}| \geqq e^{2G} - e^{G},$$

or else

$$(3) \qquad |f(x) - e^{T_n(x)}| \geqq e^{-G} - e^{-2G};$$

as $e^{2G} - e^{G} = e^{3G}(e^{-G} - e^{-2G}) > e^{-G} - e^{-2G}$, it may be asserted without distinction of alternatives that (3) is satisfied. Hence

$$g = \int_{-\pi}^{\pi} \rho(x)\, |f(x) - e^{T_n(x)}|^m\, dx \geqq (v/n)(e^{-G} - e^{-2G})^m,$$

which is inconsistent with (2) and the supposition that $g \leqq 2\gamma_n$, for all values of n from a certain point on.

Inasmuch as the $2n + 1$ coefficients of any trigonometric sum $T_n(x)$ of specified order n for which $\max |T_n(x)| < 4G$ belong to a certain bounded domain in $(2n+1)$-dimensional space, and as it is now seen that if a sequence of sums $T_n(x)$ is constructed for which the value of (1) approaches γ_n the condition $\max |T_n(x)| < 4G$ must be satisfied by all sums in the sequence from a certain point on, provided n is sufficiently large, it follows that there must be at least one limiting set of coefficients for which the integral (1) is actually equal to γ_n; the minimum problem proposed at the outset has a solution. (For completeness the possibility that $\gamma_n = 0$ requires separate notice, but the conclusion for this special case is justified with equal facility.) Furthermore (still under the supposition that n is sufficiently large) *any minimizing sum* $T_n(x)$ *satisfies the condition* $\max |T_n(x)| < 4G$. There is no assertion that the minimizing sum is uniquely determined, and no assumption to this effect will be needed in the subsequent work. The question of the existence of a minimizing sum for all values of n from the beginning, and for more general functions $f(x)$, will be considered in the next section; for the problem of convergence as n becomes infinite, with which this section is primarily concerned, any finite number of values of n may be left out of account.

In further preparation for the convergence proof an extension of Bernstein's theorem is needed, going beyond one that was presented in an earlier paper to which reference has been made. Let $f(x)$ be subject to the hypotheses already imposed, and let $T_n(x)$ be an arbitrary trigonometric sum of the n-th order; let L, h', M' be positive numbers such that

$$| f(x) - e^{T_n(x)} | \leqq L, \qquad \log h' \leqq T_n(x) \leqq \log M',$$

for all values of x. (The value $L = 0$ would be admissible but trivial.) Let h_0 be the smaller of h and h', and let $\log f(x)$ be denoted once more by $\phi(x)$. By the law of the mean,

$$f(x) - e^{T_n(x)} = e^{\phi(x)} - e^{T_n(x)} = [\phi(x) - T_n(x)]e^{\xi(x)},$$

where $\xi(x)$ has a value intermediate between $\phi(x)$ and $T_n(x)$. As $\log h_0$ is a lower bound both for $\phi(x)$ and for $T_n(x)$, it is a lower bound for $\xi(x)$, and

$$e^{\xi(x)} \geqq h_0, \qquad | \phi(x) - T_n(x) | = | f(x) - e^{T_n(x)} | e^{-\xi(x)} \leqq L/h_0.$$

Also, as it has been assumed that $| f'(x) | \leqq \lambda$,

$$| \phi'(x) | = | f'(x)/f(x) | \leqq \lambda/h_0.$$

It is possible then to apply the extension of Bernstein's theorem given in the earlier passage referred to,* with the conclusion that

$$| T'_n(x) | \leqq nL/h_0 + C\lambda/h_0,$$

where C is an absolute constant. (More specifically, the statement is true with $C = 4$, though the numerical value is not needed for present purposes.) Hence

$$| (d/dx)e^{T_n(x)} | = | T'_n(x) | e^{T_n(x)} \leqq (M'/h_0)(nL + C\lambda),$$
$$| (d/dx)[f(x) - e^{T_n(x)}] | \leqq \lambda + (M'/h_0)(nL + C\lambda).$$

The result of this calculation may be recorded in

LEMMA I. *If $f(x)$ is a positive continuous function of period 2π having a continuous first derivative subject everywhere to the condition $| f'(x) | \leqq \lambda$, and if $T_n(x)$ is a trigonometric sum of the n-th order such that*

$$\log h' \leqq T_n(x) \leqq \log M' \text{ and } | f(x) - e^{T_n(x)} | \leqq L$$

for all values of x, then

$$| (d/dx)[f(x) - e^{T_n(x)}] | \leqq C_1nL + C_2,$$

where C_1 and C_2 depend on $f(x)$ and on h' and M', but not on any other specification with regard to $T_n(x)$.

Throughout the rest of this section let $T_n(x)$, for each n, denote specifically a trigonometric sum of the n-th order for which the integral (1) has its minimum value γ_n, at least when n is large enough so that the existence of a minimizing sum is assured by the previous reasoning. It has been seen

* See *Transactions, loc. cit.*, p. 622.

that for values of n from a certain point on $|T_n(x)| < 4G$, and it will be understood that n is large enough so that this condition also is satisfied. It is to be shown that $e^{T_n(x)}$ converges uniformly toward $f(x)$ as n becomes infinite.

Let $R_n(x) \equiv f(x) - e^{T_n(x)}$, let μ_n be the maximum of $|R_n(x)|$, and let x_1 be a value of x such that $|R_n(x_1)| = \mu_n$. By application of the Lemma,

$$|R'_n(x)| \leqq C_1 n\mu_n + C_2.$$

The bounds $\log h' = -4G$, $\log M' = 4G$, on which the determination of C_1 and C_2 depends, are independent of n, and C_1 and C_2 therefore are independent of n also.

Let it be supposed temporarily that $C_2 \leqq n\mu_n$; the contrary case will be considered separately. Then

$$|R'_n(x)| \leqq (C_1+1)n\mu_n,$$
$$|R_n(x) - R_n(x_1)| \leqq (C_1+1)n\mu_n\,|x - x_1|.$$

For $|x - x_1| \leqq 1/[2(C_1+1)n]$,

$$|R_n(x) - R_n(x_1)| \leqq \tfrac{1}{2}\mu_n, \qquad |R_n(x)| \geqq \tfrac{1}{2}\mu_n.$$

Inasmuch as $\rho(x) \geqq v > 0$ everywhere,

$$\gamma_n \geqq \frac{v}{(C_1+1)n}\left(\frac{\mu_n}{2}\right)^m, \qquad \mu_n \leqq 2\left(\frac{C_1+1}{v}\right)^{1/m}(n\gamma_n)^{1/m}.$$

The supposition previously rejected, that $C_2 > n\mu_n$, would mean directly that $\mu_n < C_2/n$. In either case,

$$\mu_n \leqq 2[(C_1+1)/v]^{1/m}(n\gamma_n)^{1/m} + C_2/n.$$

Since $m \geqq 1$ it follows from (2) that $\lim\limits_{n\to\infty} n\gamma_n = 0$, $\lim\limits_{n\to\infty} \mu_n = 0$, and the uniform convergence of $R_n(x)$ toward zero is established. The conclusion is

THEOREM I. *If $f(x)$ is a positive continuous function of period 2π having a continuous derivative everywhere, if $\rho(x)$ has a positive lower bound, and if $m \geqq 1$, the sums $T_n(x)$ minimizing the integral* (1) *will be such that $e^{T_n(x)}$ converges uniformly toward $f(x)$ as n becomes infinite.*

3. **More general existence theorem.** In this section the existence of a trigonometric sum minimizing (1) is to be proved under hypotheses considerably more general than those previously admitted. The function $f(x)$, of period 2π, is assumed to be bounded and measurable, with a positive lower bound; the weight function $\rho(x)$ is of period 2π, summable, non-negative everywhere, and positive over a set of positive measure in a period; the

exponent m may have any value > 0; the order n is any positive integer or zero. The problem of uniqueness of the minimizing sum will still be left untouched.

A preliminary stage of the reasoning may be summarized in

LEMMA II. *If n is a given integer $\geqq 0$, there is for every positive ϵ a positive δ such that if $T_n(x)$ is any trigonometric sum of the n-th order having* 1 *as the maximum of its absolute value, there is a set of measure at least $2\pi - \epsilon$ in a period throughout which $|T_n(x)| \geqq \delta$.*

Suppose this were not true. Then, for some positive ϵ, there would exist a sequence of positive numbers $\delta_1, \delta_2, \cdots, \delta_k, \cdots$, approaching zero, such that for each k there is a trigonometric sum $T_{nk}(x)$, of the n-th order, having 1 as the maximum of its absolute value, and satisfying the inequality $|T_{nk}(x)| < \delta_k$ throughout a set e_k of measure greater than ϵ in the period interval $(-\pi, \pi)$. As the coefficients in the sums $T_{nk}(x)$ are bounded by the restriction $|T_{nk}(x)| \leqq 1$, the various sets of coefficients, regarded as coördinates of points in $(2n+1)$-dimensional space, must have a limit point, and there is a sum $\tau_n(x)$ of the n-th order,* still having 1 as maximum of its absolute value, uniformly approached by a sequence of the sums T_{nk}. Let $t_{n1}(x), t_{n2}(x), \cdots$, be such a sequence, the other sums T_{nk} being dismissed from further consideration. Let e be the set of points which are common to infinitely many of the corresponding sets e_k. *The measure* † *of e is at least ϵ.* If x is any point of e, $\lim_{k\to\infty} t_{nk}(x)$, which exists and is equal to $\tau_n(x)$, must be zero, by reason of the approach of the δ's to zero. So the sum $\tau_n(x)$ is required to vanish throughout a set of positive measure and to take on an extreme value ± 1, which is impossible.

An immediate corollary is that when δ has been determined for a given ϵ, in accordance with the terms of the Lemma, then *if $T_n(x)$ is any trigonometric sum of the n-th order whatever, and if $|T_n(x)|$ attains a value as large as H,* whatever the value of H may be, *then $|T_n(x)| \geqq H\delta$ throughout a set of measure at least $2\pi - \epsilon$ in a period.*

To return to the problem of minimizing the integral (1), let $h > 0$ be a lower bound and M an upper bound for $f(x)$. Let $T_n(x)$ be an arbitrary trigonometric sum of the n-th order, let

$$g_1 = \int_{-\pi}^{\pi} \rho(x) \, | f(x) - e^{T_n(x)} |^m \, dx,$$

---

* Whenever reference is made to sums of the n-th order, the words are understood to mean of the n-th order *at most.*

† See e. g. de la Vallée Poussin, *Intégrales de Lebesgue,* Paris, 1916, pp. 8-9, 26-27.

and let
$$g_2 = \int_{-\pi}^{\pi} \rho(x)\,[f(x) - h]^m\,dx;$$
the quantity in brackets in the last integral is non-negative, and the constant $h = e^{\log h}$ may be regarded as a function of the form $e^{T_n(x)}$ for any value of $n \geqq 0$.

Let h_1 be a positive number less than h: $0 < h_1 < h$. Let E_1 be the set of points in $(-\pi, \pi)$ (if any) at which $e^{T_n(x)} \leqq h - h_1$, let E_2 be the set where $e^{T_n(x)} \geqq 2M - h + h_1$, let $E = E_1 + E_2$, and let CE be the set complementary to E, on which
$$h - h_1 < e^{T_n(x)} < 2M - h + h_1.$$
Let $r_1(x) = f(x) - e^{T_n(x)}$, $r_2(x) = f(x) - h \geqq 0$. Any point of E_1,
$$|\,r_1(x)\,| - r_2(x) = r_1(x) - r_2(x) = h - e^{T_n(x)} \geqq h_1.$$
At any point of E_2, as $f(x) \leqq M$,
$$|\,r_1(x)\,| \geqq M - h + h_1,$$
and as $r_2(x) \leqq M - h$ everywhere,
$$|\,r_1(x)\,| - r_2(x) \geqq h_1,$$
for x in E_2 as well as for x in E_1. If $m > 1$, inasmuch as $r_2 \geqq 0$, the last inequality implies that
$$|\,r_1\,|^m - r_2{}^m \geqq (r_2 + h_1)^m - r_2{}^m \geqq h_1{}^m,$$
the difference $(r_2 + h_1)^m - r_2{}^m$ being smaller for $r_2 = 0$ than for any positive value of r_2; if $m < 1$, inasmuch as $r_2 \leqq M - h$, the corresponding inference is that
$$|\,r_1\,|^m - r_2{}^m \geqq (r_2 + h_1)^m - r_2{}^m \geqq (M - h + h_1)^m - (M - h)^m,$$
the difference $(r_2 + h_1)^m - r_2{}^m$ now being less for $r_2 = M - h$ than for any smaller non-negative value of r_2. If D_1 denotes $h_1{}^m$ when $m > 1$, $(M - h + h_1)^m - (M - h)^m$ when $m < 1$, and the common value h_1 to which both expressions reduce when $m = 1$, then $|\,r_1\,|^m - r_2{}^m \geqq D_1$ throughout E, and
$$J_1 = \int_E \rho(x)\,(|\,r_1\,|^m - r_2{}^m)\,dx \geqq D_1 \int_E \rho(x)\,dx.$$
With regard to points of CE it can be said that
$$|\,r_1\,|^m - r_2{}^m \geqq -r_2{}^m \geqq -(M - h)^m,$$
and if $(M - h)^m$ is denoted by D_2,
$$J_2 = \int_{CE} \rho(x)\,(|\,r_1\,|^m - r_2{}^m)\,dx \geqq -D_2 \int_{CE} \rho(x)\,dx.$$

So

$$g_1 - g_2 = J_1 + J_2 \geqq D_1 \int_E \rho(x)\,dx - D_2 \int_{CE} \rho(x)\,dx.$$

By the absolute continuity of $\int \rho(x)\,dx$, the integral over CE can be brought arbitrarily near to zero, and the integral over E arbitrarily near to the integral over an entire period, if the measure of CE, denoted by mCE, can be made sufficiently small. In particular, if

$$I = \int_{-\pi}^{\pi} \rho(x)\,dx,$$

and if η is a positive number less than $D_1 I/(D_1 + D_2)$, there will be a positive ϵ such that

$$\int_{CE} \rho(x)\,dx < \eta$$

if $mCE \leqq \epsilon$, and then it will follow further that

$$\int_E \rho(x)\,dx > I - \eta,$$

$$g_1 - g_2 > D_1(I - \eta) - D_2\eta > 0.$$

Let δ be the quantity associated with this ϵ by Lemma II. Let H_1 be the larger of the numbers $|\log(h - h_1)|$, $\log(2M - h + h_1)$. If $|T_n(x)| \geqq H_1$, it will follow that

$$T_n(x) \leqq -H_1 \leqq -|\log(h - h_1)| \leqq \log(h - h_1), \quad e^{T_n(x)} \leqq h - h_1,$$

or else

$$T_n(x) \geqq H_1 \geqq \log(2M - h + h_1), \quad e^{T_n(x)} \geqq 2M - h + h_1.$$

That is to say, any x for which $|T_n(x)| \geqq H_1$ belongs to the set E. By the Lemma, as interpreted through its corollary, if $|T_n(x)|$ attains anywhere a value as large as H_1/δ, the measure of E will be at least $2\pi - \epsilon$, the measure of CE will be not more than ϵ, and according to the preceding paragraph $g_1 - g_2$ will be positive.

This means that all sums $T_n(x)$ for which $g_1 \leqq g_2$ are such that $|T_n(x)| < H_1/\delta$ for all values of x, and the coefficients are thereby required to belong to a bounded domain. If the greatest lower bound γ_n of the integral (1) is equal to g_2, the constant $h = e^{\log h}$ is itself a minimizing sum; if not, then $\gamma_n < g_2$, any sequence of sums $T_n(x)$ for which the value of the integral approaches γ_n will have coefficients belonging to the bounded domain just mentioned from a certain point on, and there will necessarily be a limiting set of coefficients and a corresponding sum $T_n(x)$ for which the value γ_n is attained. Thus the existence of a minimizing sum is assured:

THEOREM II. *Under the conditions stated at the beginning of the section, there will be for each* $n \geqq 0$ *at least one sum* $T_n(x)$ *for which the value of the integral* (1) *is a minimum.*

4. **Polynomial approximation.** A corresponding problem of polynomial approximation over a finite interval, which may without loss of generality be taken as that from -1 to 1, relates to the minimizing of the integral

$$(4) \qquad \int_{-1}^{1} \rho(x) \mid f(x) - e^{P_n(x)} \mid^m dx,$$

in which $P_n(x)$ is a polynomial of the n-th degree (at most).

The existence proof of the preceding section can be adapted to this case without difficulty; under hypotheses similar in generality to those formulated at the beginning of Section 3 there exists for each value of n a minimizing polynomial, which may or may not be uniquely determined.

The circumstances of the proof of convergence are materially changed, though not to the extent of obliterating the analogy, by the fact that Bernstein's theorem and its generalization are less simple for polynomials than for trigonometric sums.

The hypothesis with regard to $f(x)$ (in addition to the requirement that it take on only positive values) will be once more that it have a continuous derivative, in the present case for $-1 \leqq x \leqq 1$. The exponent m, however, will be restricted to values $\geqq 2$. It will be supposed again that $\rho(x)$ has a positive lower bound. It can be shown by appropriate modification of the previous argument that the minimizing polynomials are uniformly bounded for all values of n. The question which then calls for special attention is the adaptation of Bernstein's theorem.

Let $f(x)$ be continuous and positive for $-1 \leqq x \leqq 1$, having $h > 0$ and M as its minimum and maximum values, and let $f'(x)$ be defined and continuous throughout the interval, with λ as an upper bound for its absolute value. Let $P_n(x)$ be an arbitrary polynomial of the n-th degree, and let

$$| f(x) - e^{P_n(x)} | \leqq L, \qquad \log h' \leqq P_n(x) \leqq \log M',$$

for $-1 \leqq x \leqq 1$.

Let $x = \cos \theta$. Then $f(x) = f(\cos \theta) = F(\theta)$ is a periodic function of θ, defined and continuous for all real values of the variable. Its minimum and maximum values are those of $f(x)$, and it has furthermore a continuous derivative $F'(\theta) = -f'(x) \sin \theta$, subject to the inequality $| F'(\theta) | \leqq \lambda$. Also, $P_n(x) = P_n(\cos \theta)$ is a trigonometric sum of the n-th order in θ, which may be represented by $T_n(\theta)$. The bounds of $T_n(\theta)$ are those of $P_n(x)$, and

$$| F(\theta) - e^{T_n(\theta)} | = | f(x) - e^{P_n(x)} | \leqq L$$

for all values of θ. Lemma I is therefore directly applicable, with θ as independent variable, to the effect that

$$| (d/d\theta)[F(\theta) - e^{T_n(\theta)}] | \leqq C_1 nL + C_2.$$

For differentiation with respect to x, as

$$\begin{aligned}(d/dx)[f(x) - e^{P_n(x)}] &= (d/dx)[F(\theta) - e^{T_n(\theta)}] \\ &= (d/d\theta)[F(\theta) - e^{T_n(\theta)}](d\theta/dx),\end{aligned}$$

this means that

$$| (d/dx)[f(x) - e^{P_n(x)}] | \leqq \frac{C_1 nL + C_2}{(1-x^2)^{1/2}}.$$

In formal statement:

LEMMA III. *If $f(x)$ is a positive continuous function for $-1 \leqq x \leqq 1$, having throughout the interval a continuous first derivative subject to the condition $| f'(x) | \leqq \lambda$, and if $P_n(x)$ is a polynomial of the n-th degree such that $\log h' \leqq P_n(x) \leqq \log M'$ and $| f(x) - e^{P_n(x)} | \leqq L$ for $-1 \leqq x \leqq 1$, then for $-1 < x < 1$,*

$$| (d/dx)[f(x) - e^{P_n(x)}] | \leqq \frac{C_1 nL + C_2}{(1-x^2)^{1/2}},$$

where C_1 and C_2 depend on $f(x)$ and on h' and M', but not on any other specification with regard to $P_n(x)$.

Repetition of an argument used elsewhere * in connection with the ordinary form of Bernstein's theorem leads directly to the

COROLLARY. *If x_1 and x_2 are any two numbers of the closed interval $(-1, 1)$ differing by not more than unity, and if $f(x) - e^{P_n(x)} \equiv R_n(x)$,*

$$| R_n(x_2) - R_n(x_1) | \leqq 2(C_1 nL + C_2) | x_2 - x_1 |^{1/2}.$$

With obvious readjustments, and in particular with replacement of an interval whose length is of the order of $1/n$ by an interval whose length is of the order of $1/n^2$, the reasoning which gave a proof of Theorem I now shows that if $P_n(x)$ is for each value of n a polynomial minimizing the integral (4), and if $m \geqq 2$, *$e^{P_n(x)}$ converges uniformly toward $f(x)$ for $-1 \leqq x \leqq 1$ as n becomes infinite.*

THE UNIVERSITY OF MINNESOTA,
MINNEAPOLIS.

* See the writer's *Colloquium*, previously cited, pp. 93-94.

CERTAIN IRREGULAR NON-HOMOGENEOUS LINEAR DIFFERENCE EQUATIONS.*

By David Moskovitz.

1. *Introduction.* Among the first to obtain theorems of existence for the solutions of a linear difference equation was Birkhoff,† who treated the so-called "regular" case in which the characteristic equation of the difference equation (or system of equations) has no infinite, zero, or multiple roots. Other writers who contributed to the study of this problem at about the same time were Nörlund ‡ and Carmichael.§ In 1913, Williams,¶ employing the methods of Birkhoff, examined the non-homogeneous equation under the condition that the associated homogeneous equation is regular, obtained "principal" solutions, and found their asymptotic forms.

Comparatively little was accomplished in studying the "irregular" cases of the homogeneous equation until Adams ‖ obtained results for certain classes of the irregular cases. The most recent contribution to the irregular case problem for the homogeneous equation is by Birkhoff,** who has given formal solutions in all possible irregular cases; a treatment of the corresponding analytic theory by Birkhoff and Trjitzinsky, is expected to appear soon.††

---

* Presented to the American Mathematical Society, November 25, 1932.

† G. D. Birkhoff, "General theory of linear difference equations," *Transactions of the American Mathematical Society*, Vol. 12 (1911), p. 243. (This paper will hereafter be referred to as *B.*)

‡ N. E. Nörlund, "Sur les Équations aux Différences Finies," *Comptes Rendus*, Vol. 149 (1909), p. 841.

§ R. D. Carmichael, "Linear difference equations and their analytic solutions," *Transactions of the American Mathematical Society*, Vol. 12 (1911), p. 99.

¶ K. P. Williams, "The solutions of non-homogeneous linear difference equations and their asymptotic form," *Transactions of the American Mathematical Society*, Vol. 14 (1913), p. 209. (This paper will hereafter be referred to as *W.*)

‖ C. R. Adams, "On the irregular cases of the linear ordinary difference equation," *Transactions of the American Mathematical Society*, Vol. 30 (1928), p. 507. (This paper will hereafter be referred to as *A.*)

** G. D. Birkhoff, "Formal theory of irregular linear difference equations," *Acta Mathematica*, Vol. 54 (1930), p. 205.

†† This paper has appeared while the present paper was in press: G. D. Birkhoff and W. J. Trjitzinsky, "Analytic theory of singular difference equations," *Acta Mathematica*, Vol. 60 (1933), pp. 1-89.

The purpose of the present paper is to establish the existence of analytic solutions of the non-homogeneous linear difference equation

$$(1.1)\qquad \sum_{k=0}^{n} a_k(x)y(x+n-k) = b(x),$$

in which the functions $a_k(x)$ and $b(x)$ are assumed to be rational,* and for which the associated homogeneous equation

$$(1.2)\qquad \sum_{k=0}^{n} a_k(x)y(x+n-k) = 0$$

belongs to the class of irregular cases which are called class 2a in A (p. 513), and to obtain the asymptotic forms of these solutions.

Two methods are employed, one analogous to that used in W in studying the non-homogeneous equation whose associated homogeneous equation is regular; by this method the results obtained by Williams are extended to apply to the class of equations here studied, and certain of his results are amplified.

In § 2, we give the results from A which we use in our development. In § 3, we give symbolic solutions of the problem of summation. In § 4, we obtain a formal power series solution of equation (1.1). §§ 5, 6, and 7 are devoted to the existence and asymptotic properties of a first analytic solution of (1.1) under the assumption that none of the segments of the broken line L of A (p. 511) have positive slopes and that the absolute values of the roots of the characteristic equation associated with the horizontal segment of L are greater than one. We show in § 7 that the first solution is asymptotically represented by the formal power series uniformly throughout a region which is not too close to the negative axis of reals. We show that the first solution is the only analytic solution of (1.1) which is represented by the formal power series solution in any domain which has in it at least one horizontal line extending to infinity to the right. These results are also applicable to the problem treated in W, and attention is called to this fact.

In § 8, we obtain a second analytic solution which is asymptotically represented by the formal power series in the left half plane.

Under the assumption that none of the slopes of the segments of the broken line L are negative and the absolute values of the roots of the characteristic equation associated with the horizontal segment of L are less than

* The results which we shall establish are valid, with slight modification, in the case where the functions $a_k(x)$ and $b(x)$ are assumed to be of a rational character only at infinity. The only modification necessary is that the singularities of the solutions obtained may be other than poles.

one, we obtain similar results in which the rôle of right and left are interchanged. These results are given in § 9.

In the more general case in which some of the slopes of the segments of the broken line L are positive and some negative, our results by methods used thus far are not as complete as in the special cases treated previously. By a second method which appears in § 10, we obtain results for our problem where the sole restriction concerning the characteristic equation associated with the horizontal segment of L is that it does not have unity as a root. These results are at once applicable to the case $\mu = 0$ in W (pp. 234 *et seq.*), whose treatment is open to objection. Our second method consists in transforming the non-homogeneous equation to a homogeneous equation of order $(n+1)$, which is shown to belong to the same irregular case (Class 2a of A, p. 513), as the homogeneous equation (1.2). The results in A concerning the solutions of homogeneous equations in this irregular case may be applied at once.

2. *Existence theorems for the homogeneous equation.* Adams has shown in A (p. 513) that the equation (1.2) has n formal series solutions

$$(2.1)\quad s_k(x) = (x/e)^{\mu_k x}\rho_k^{\,x}x^{r_k}(1 + s_{k,1}/x + s_{k,2}/x^2 + \cdots) \quad [k = 1, 2, \cdots, n],$$

which are in general divergent. The numbers μ_k are the negatives of the slopes of the segments of the broken line L of Figure 1 of A (p. 511), and the numbers ρ_k are the roots of the several characteristic equations, one of which is associated with each segment of the broken line L. The characteristic equation associated with the segment L_{cd} of L for which $\mu_c \neq \mu_{c+1} = \mu_{c+2} = \cdots = \mu_d \neq \mu_{d+1}$ is given by $\sum_{k=c}^{d} a_{k,0}\rho^{d-k} = 0$.

The equation (1.2) has two sets of solutions $h_1(x), h_2(x), \cdots, h_n(x)$ $[g_1(x), g_2(x), \cdots, g_n(x)]$, having the properties that they are analytic throughout the finite plane except for poles within a left [right] P-region,* and $h_k(x)$ $[g_k(x)]$ is asymptotically represented by $s_k(x)$ in the sector $-\pi/2 < \arg x < \pi/2$ $[\pi/2 < \arg x < 3\pi/2]$.

If we let $s_k^{(t)}(x)$ denote the sum of the first $(t+1)$ terms of $s_k(x)$, then the following equations express the asymptotic properties † of the above solutions of (1.2).

* We shall mean by a right [left] P-region of radius R, that part of the plane which is bounded by a semi-circle of radius R about the origin lying entirely to the left [right] of the imaginary axis, and two half-lines which are tangent to the semi-circle and extend to infinity parallel to the positive [negative] axis of reals.

† In order to facilitate the writing we shall use $C(x)$, either with or without a subscript to denote generically a function which is uniformly bounded for all x in the region shown.

$$(2.2)\quad \begin{aligned} h_k(x) &= s_k^{(t)}(x)[1 + C(x)/x^{t+1}] \qquad (-\pi/2 < \arg x < \pi/2),\\ g_k(x) &= s_k^{(t)}(x)[1 + C(x)/x^{t+1}] \qquad (\pi/2 < \arg x < 3\pi/2). \end{aligned}$$

We shall also make use of the two sets of "intermediate" solutions of the equation (1.2) denoted by $h'_1(x)$, $h'_2(x), \cdots, h'_n(x)$ and $g'_1(x)$, $g'_2(x), \cdots, g'_n(x)$, which are defined and analytic only in the finite part of the plane for which $|v| > R$, $[x = u + (-1)^{1/2} v]$, and which have the following asymptotic properties *

$$(2.3)\quad \begin{aligned} h'_k(x) &= s_k^{(t)}(x)[1 + C(x)/x^{t+1}] \qquad (u \geqq u_0,\ |v| > R),\\ &= s_k^{(t)}(x)\left[1 + \frac{C(x)}{v^{t+1}}\left(\frac{x}{v}\right)^{\delta}\right] \qquad (u \leqq u_0,\ |v| > R),\\ g'_k(x) &= s_k^{(t)}(x)[1 + C(x)/x^{t+1}] \qquad (u \leqq u_0,\ |v| > R),\\ &= s_k^{(t)}(x)\left[1 + \frac{C(x)}{v^{t+1}}\left(\frac{x}{v}\right)^{\delta}\right] \qquad (u \geqq u_0,\ |v| > R), \end{aligned}$$

where u_0 is any real number and δ is a non-negative constant which is independent of t; for a first order equation $\delta = 0$.

The following relations between the solutions (2.2) and (2.3) are also useful

$$(2.4)\quad \begin{aligned} h_k(x) &= \theta_{n,k}(x)h'_n(x) + \theta_{n-1,k}(x)h'_{n-1}(x) + \cdots\\ &\qquad + \theta_{k+1,k}(x)h'_{k+1}(x) + h'_k(x),\\ g_k(x) &= \phi_{1,k}(x)g'_1(x) + \phi_{2,k}(x)h'_2(x) + \cdots\\ &\qquad + \phi_{k-1,k}(x)g'_{k-1}(x) + g'_k(x), \end{aligned}$$

in which the functions $\theta_{ij}(x)$ and $\phi_{ij}(x)$ $(i, j = 1, 2, \cdots, n)$ are periodic of period one, and are analytic in any finite region lying in the part of the plane defined by $|v| > R$.

3. *Symbolic solutions.* Symbolic solutions of the non-homogeneous equation (1.1) can be derived from those of the associated homogeneous equation (1.2) by a method analogous to that of variation of parameters used in solving a non-homogeneous linear differential equation. Let $y_k(x)$ $[k = 1, 2, \cdots, n]$ be a fundamental set † of solutions of the equation (1.2), and put

---

* The results concerning asymptotic form with respect to x, as obtained in *B* and *A*, may be expressed in the form of the first and third of the equations of (2.3); Birkhoff's results concerning asymptotic form with respect to v are stated not quite precisely, and the same criticism can be made of Adams' paper. The result was correctly given by Nörlund in his *Leçons sur les Équations Linéaires aux Differences Finies*, Paris (1929), pp. 130-152. The correct asymptotic relations are expressed here by the second and fourth equations of (2.3).

† Let $y_1(x)$, $y_2(x), \cdots, y_n(x)$ be solutions of (1.2) which are analytic in a

$$y(x)=\sum_{k=1}^{n} w_k(x)y_k(x). \tag{3.1}$$

The function $y(x)$ will be a solution of (1.1) if the differences of the functions $w_k(x)$ satisfy *

$$\Delta w_k(x)=(-1)^{n+k}\frac{b(x)Y_k(x)}{a_0(x)Y(x)}=\theta_k(x) \qquad [k=1,2,\cdots,n], \tag{3.2}$$

in which † $Y(x)=\mathrm{Det}\,[y_j(x+i)]$ and $Y_k(x)=\mathrm{Det}(k)[y_j(x+i)]$.

The equation (3.2) has two formal series solutions

$$\begin{aligned} w_k(x)&=-\sum_{m=0}^{\infty}\theta_k(x+m),\\ w_k(x)&=\sum_{m=1}^{\infty}\theta_k(x-m), \end{aligned} \tag{3.3}$$

which we term the "symbolic series solutions" to the right and to the left, respectively, and which if they converge will be actual solutions of (3.2). A "symbolic contour integral solution" of (3.2) is given by

$$w_k(x)=\int_{\Gamma}\frac{\theta_k(z)\,dz}{e^{2\pi(x-z)(-1)^{1/2}}-1}, \tag{3.4}$$

where Γ is the contour $\infty AB\infty$ of Figure 1 of W (p. 215).

In seeking analytic solutions of (1.1) our problem is to determine whether we can choose a fundamental set of solutions of (1.2) so that one or more of the symbolic solutions of (3.2) yield analytic solutions of that equation in which case (3.1) yields an analytic solution of (1.1).

4. *Formal power series solutions.* If the numbers μ_k have the following distribution of signs

$$\begin{gathered}\mu_k>0\ [k=1,2,\cdots,\alpha];\quad \mu_k=0\ [k=\alpha+1,\ \alpha+2,\cdots,\beta];\\ \mu_k<0\ [k=\beta+1,\cdots,n],\end{gathered} \tag{4.1}$$

the expansions of the coefficient functions $a_k(x)$ can be written in the following way:

---

region Q; if there exist n periodic functions $\pi_k(x)$ $[k=1,2,\cdots,n]$ which are analytic in Q and which are not all identically zero, such that the expression $\pi_1(x)y_1(x)+\pi_2(x)y_2(x)+\cdots+\pi_n(x)y_n(x)$ vanishes identically, the analytic solutions $y_k(x)$ $[k=1,2,\cdots,n]$ are said to be linearly dependent in Q; otherwise they are linearly independent and form a fundamental set of solutions in the region Q.

* For details, see, for example, Batchelder, *An introduction to Linear Difference Equations*, Harvard University Press, 1927, p. 13.

† We shall use the notation Det $[a_{ij}]$ to represent the n-th order determinant whose element in the i-th row and the j-th column is a_{ij}; and Det $(k)[a_{ij}]$ will denote the minor of the element in the last row and the k-th column of Det $[a_{ij}]$.

$$a_k(x) = \frac{1}{x^{\mu_{k+1}+\mu_{k+2}+\cdots+\mu_\alpha}}(a_{k,0} + a_{k,1}/x + a_{k,2}/x^2 + \cdots) \qquad [k = 0, 1, 2, \cdots, \alpha - 1],$$

$$(4.2) \quad a_k(x) = a_{k,0} + a_{k,1}/x + a_{k,2}/x^2 + \cdots \qquad [k = \alpha, \alpha + 1, \cdots, \beta],$$

$$a_k(x) = \frac{1}{x^{-\mu_{\beta+1}-\mu_{\beta+2}-\cdots-\mu_k}}(a_{k,0} + a_{k,1}/x + a_{k,2}/x^2 + \cdots) \qquad [k = \beta + 1, \beta + 2, \cdots, n],$$

in which, at least those numbers $a_{k,0}$ corresponding to the values of k at which the broken line L changes its slope are different from zero. The expansions (4.2) as well as the following

$$(4.3) \qquad b(x) = x^\lambda(b_0 + b_1/x + b_2/x^2 + \cdots),$$

are valid for $|x| > \bar{R}$.

We can find a power series which formally satisfies (1.1) by substituting the series

$$\bar{y}(x) = x^m(y_0 + y_1/x + y_2/x^2 + \cdots)$$

into the equation (1.1). Replacing the functions $b(x)$ and $a_k(x)$ by their expansions as given in (4.2) and (4.3), we find the highest degree term on the left side of the resulting equation to be x^m, while on the right side it is x^λ. Set $m = \lambda$, and equate coefficients of like powers of x, and we find that $y_0 = b_0/\sum_{k=\alpha}^{\beta} a_{k,0}$. The characteristic equation associated with the horizontal segment of the broken line L is $\sum_{k=\alpha}^{\beta} a_{k,0}\rho^{\beta-k} = 0$. We assume that this equation does not have the root $\rho = 1$, and hence $\sum_{k=\alpha}^{\beta} a_{k,0} \neq 0$, and y_0 is uniquely determined. The coefficients $y_1, y_2, \cdots$ can be determined successively and uniquely since

$$y_s = [b_s/\sum_{k=\alpha}^{\beta} a_{k,0}] + \text{terms in } y_0, y_1, \cdots, y_{s-1} \qquad [s = 1, 2, 3, \cdots].$$

When the broken line L has no horizontal segment, and the numbers μ_k have the following distribution of signs

$$\mu_k > 0 \quad [k = 1, 2, \cdots, \alpha]; \qquad \mu_k < 0 \quad [k = \alpha + 1, \alpha + 2, \cdots, n],$$

the leading coefficient in the formal power series solution is given by $y_0 = b_0/a_{\alpha,0}$, which is uniquely determined since $a_{\alpha,0} \neq 0$.

5. *The first analytic solution.* We shall now obtain an analytic solution of the equation (1.1) under the assumption that

(5.1) $\mu_k > 0$ $[k = 1, 2, \cdots, \gamma]$; $\mu_k = 0$, $|\rho_k| > 1$ $[k = \gamma + 1, \gamma + 2, \cdots, n]$.

Use the functions $h_k(x)$ $[k = 1, 2, \cdots, n]$ as the solutions of (1.2), and the symbolic series solution to the right of equation (3.2). We thus have

$$(5.2) \qquad w_k(x) = -\sum_{m=0}^{\infty} T_k(x+m) \qquad [k = 1, 2, \cdots, n],$$

where

$$T_k(x) = (-1)^{n+k} \frac{b(x) H_k(x)}{a_0(x) H(x)},$$

and in which $H(x) = \operatorname{Det}[h_j(x+i)]$ and $H_k(x) = \operatorname{Det}(k)[h_j(x+i)]$.

All the singularities of $T_k(x)$ can be enclosed in a left P-region of sufficiently large radius R;* let P denote this region; we shall show that the series (5.2) are uniformly convergent in any finite region exterior to P. Let V be that part of the sector $-\pi/2 < \arg x < \pi/2$ which is exterior to P. From (2.2), we have for all x in V,

$$(5.3) \quad h_j(x+i) = \left(\frac{x+i}{e}\right)^{\mu_j(x+i)} \rho_j^{x+i} (x+i)^{r_j}$$
$$\times \left[1 + \frac{s_{j,1}}{x+i} + \cdots + \frac{s_{j,t}}{(x+i)^t}\right]\left[1 + \frac{C_j(x+i)}{(x+i)^{t+1}}\right]$$
$$= (x/e)^{\mu_j x} \rho_j^x x^{r_j} (\rho_j x^{\mu_j})^i C_{ij}(x),$$

where

$$C_{ij}(x) = (1 + i/x)^{\mu_j(x+i)+r_j} e^{-\mu_j i} \left[1 + \frac{s_{j,1}}{x+i} + \cdots + \frac{s_{j,t}}{(x+i)^t}\right]$$
$$\times \left[1 + \frac{C_j(x+i)}{(x+i)^{t+1}}\right] = 1 + O(1/x)$$

is uniformly bounded in V. The elements of the j-th column of $H(x)$ have the common factor $(x/e)^{\mu_j x} \rho_j^x x^{r_j}$, and therefore we may write

$$(5.4) \qquad H(x) = (x/e)^{(\mu_1+\mu_2+\cdots+\mu_n)x} (\rho_1 \rho_2 \cdots \rho_n)^x x^{r_1+r_2+\cdots+r_n} J(x),$$

where $J(x) = \operatorname{Det}[(\rho_j x^{\mu_j})^i C_{ij}(x)]$. The expansion of the determinant $J(x)$ may be written in the form $J(x) = \sum_{m=1}^{n!} \Lambda_m \Omega_m(x) x^{\Gamma_m}$, where Λ_m is the product of n factors of the form $\rho_j{}^i$, $\Omega_m(x)$ is the product of n factors of the form $C_{ij}(x)$, and Γ_m is the sum of n terms of the form $i\mu_j$ $(i, j = 1, 2, \cdots, n)$. Since $\mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_n$, it is apparent that the largest number of the set Γ_m is the number

$$(5.5) \qquad E = n\mu_1 + (n-1)\mu_2$$
$$+ \cdots + (n-j+1)\mu_j + \cdots + \mu_n = \sum_{j=1}^{n} (n-j+1)\mu_j.$$

---

* If the singularities of $T_k(x)$ can be enclosed in a left P-region of radius less than n, choose $R > n$; also choose $R > \bar{R}$.

When the following relation is satisfied

$$\mu_1 = \mu_2 = \cdots = \mu_\alpha \neq \mu_{\alpha+1}$$
$$= \mu_{\alpha+2} = \cdots = \mu_\beta \neq \cdots \neq \mu_{\gamma+1} = \mu_{\gamma+2} = \cdots = \mu_n,$$

the number Λ_m which is the coefficient of x^E is given by the determinant

$$A = \begin{vmatrix} & 0 & & \cdot\; A^{n-\gamma,\, n-\gamma-1, \ldots, 1}_{\gamma+1, \ldots, n} \\ & & \cdot & \\ & A^{n-\alpha, \ldots, n-\beta+1}_{\alpha+1, \alpha+2, \ldots, \beta} \;\cdot & & \\ A^{n,\, n-1, \ldots, n-\alpha+1}_{1, 2, \ldots, \alpha} & & & 0 \end{vmatrix},$$

in which

$$A^{l, m, \ldots, n}_{r, s, \ldots, t} = \begin{vmatrix} \rho_r^{\,n} & \rho_s^{\,n} & \cdots & \rho_t^{\,n} \\ \cdot & \cdot & \cdot & \cdot \\ \rho_r^{\,m} & \rho_s^{\,m} & \cdots & \rho_t^{\,m} \\ \rho_r^{\,l} & \rho_s^{\,l} & \cdots & \rho_t^{\,l} \end{vmatrix}.$$

The determinant A can be written as the product of factors of the form $A^{l, m, \ldots, n}_{r, s, \ldots, t}$ in which the number of factors is equal to the number of segments in the broken line L; each of these factors are of Vandermondean type in which the ρ's are distinct and different from zero, and hence A is different from zero. Therefore, we may write $J(x) = Ax^E\,[1 + \bar{M}(x)]$, where $\bar{M}(x)$ is uniformly bounded in V, and $\lim\limits_{x\to\infty \text{ in } V} \bar{M}(x) = 0$. Similarly, we can show that

$$H_k(x) = \frac{(x/e)^{(\mu_1+\mu_2+\cdots+\mu_n)x}\,(\rho_1\rho_2\cdots\rho_n)^x\, x^{r_1+r_2+\cdots+r_n}}{(x/e)^{\mu_k x}\,\rho_k^{\,x}\, x^{r_k}}\, J_k(x),$$

where $\quad J_k(x) = \operatorname{Det}(k)\,[(\rho_j x^{\mu_j})^i\, C_{ij}(x)] = A_k x^{E_k}[1 + \bar{M}_k(x)],$

in which $\qquad E_k = \sum\limits_{j=1}^{k-1} (n-j)\mu_j + \sum\limits_{j=k+1}^{n} (n-j+1)\mu_j$

and

$$A_k = \begin{vmatrix} & & & & \cdot\; A^{n-\gamma,\, n-\gamma-1, \ldots, 1}_{\gamma+1, \gamma+2, \ldots, n} \\ & 0 & & \cdot & \\ & & \cdot\; A^{n-p-1,\, n-p-2, \ldots, n-q+1}_{p+1, p+2, \ldots, k-1, k+1, \ldots, q} \;\cdot & & \\ & A^{n-\alpha-1, \ldots, n-\beta}_{\alpha+1, \ldots, \beta} \;\cdot & & & 0 \\ A^{n-1,\, n-2, \ldots, n-\alpha}_{1, 2, \ldots, \alpha} & & & & \end{vmatrix}$$

and $\bar{M}_k(x)$ is uniformly bounded in V, and $\lim\limits_{x\to\infty \text{ in } V} \bar{M}_k(x) = 0.$

Therefore

$$T_k(x) = (-1)^{n+k}\, \frac{x^{\lambda+\mu_1+\mu_2+\cdots+\mu_n}}{(x/e)^{\mu_k x}\,\rho_k^{\,x}\, x^{r_k+E-E_k}}\, d_k[1 + M_k(x)],$$

where $M_k(x)$ is uniformly bounded in V and $\lim_{x\to\infty \text{ in } V} M_k(x) = 0$; and

$$d_k = b_0 A_k / a_{0,0} A \qquad [k = 1, 2, \cdots, n].$$

The $(m+1)$-th term of the series (5.2) can be written in the form

$$\frac{(-1)^{n+k+1} d_k[1 + M_k(x+m)]}{(x+m)^{\mu_k(x+m)+c_k} (\rho_k e^{-\mu_k})^{x+m}},$$

where $c_k = r_k + E - E_k - (\lambda + \mu_1 + \mu_2 + \cdots + \mu_n)$. Let Q be any finite region exterior to P; in this region the series (5.2) can be shown to be dominated by

$$M \sum_{m=0}^{\infty} \frac{1}{|\rho_k e^{-\mu_k}|^m \cdot m^{c'_k + \mu_k m}}, \tag{5.6}$$

where c'_k = real part of c_k and M is a constant which depends only on the region Q. The series (5.6) converges when $\mu_k > 0$, and also when $\mu_k = 0$, $|\rho_k| > 1$; therefore the series (5.2) converges uniformly for all x in Q. Consequently the series (5.2) represents an analytic function provided that x is finite and exterior to P, since each of the terms of (5.2) is analytic at each such point x. If x is within P, the first few terms of (5.2) may have poles while the remaining part of the series converges uniformly; accordingly $w_k(x)$ may have poles within P.

When the values of $w_k(x)$ thus obtained are substituted into $h(x) = \sum_{k=1}^{n} w_k(x) h_k(x)$, we have a solution of (1.1) which is analytic in the finite part of the plane except perhaps for poles within P. The solution thus obtained can be extended to the left by means of the equation (1.1) itself from which we find

$$h(x) = \frac{b(x)}{a_n(x)} - \sum_{k=0}^{n-1} \frac{a_k(x)}{a_n(x)} h(x+n-k),$$

which defines $h(x)$ at all finite points which are within P but which are not poles of $b(x)$ nor $a_k(x)\,[k = 0, 1, 2, \cdots, n-1]$, nor zeros of $a_n(x)$, nor points congruent to all these on the left.

6. *A property of the solution obtained by using the symbolic series.* Solutions of the equation (1.1) obtained by the method of the preceding section have an important property which is stated in the following lemma, which we shall prove in this section.

LEMMA. *Let Q' and Q'' be two regions in which it is possible to iterate to the right,* which have in common the region Q. Let $h_1(x), h_2(x), \cdots,$*

* Iteration to the right [left] is possible in a region if it has the property that if x is any point of the region, then also the points $x+1, x+2, \cdots, x+k, \cdots$ $[x-1, x-2, \cdots, x-k, \cdots]$ are in the region.

$h_n(x)$ be a fundamental set of solutions of the equation (1.2) in the region Q', and let $h(x) = \sum_{k=1}^{n} w_k(x) h_k(x)$ be a solution of (1.1) which is analytic in Q', and where the functions $w_k(x)$ are analytic in Q' and defined by

$$w_k(x) = (-1)^{n+k+1} \sum_{m=0}^{\infty} \frac{b(x+m) H_k(x+m)}{a_0(x+m) H(x+m)} \qquad [k = 1, 2, \cdots, n],$$

which are assumed to be convergent in Q'. Let $y_1(x), y_2(x), \cdots, y_n(x)$ be another set of solutions of (1.2) which form a fundamental set of solutions in the region Q''. Then the series

$$w'_k(x) = (-1)^{n+k+1} \sum_{m=0}^{\infty} \frac{b(x+m) Y_k(x+m)}{a_0(x+m) Y(x+m)} \qquad [k = 1, 2, \cdots, n],$$

converge in Q, and the function $y(x)$ defined by $y(x) = \sum_{k=1}^{n} w'_k(x) y_k(x)$ is identical with $h(x)$ in the region Q.

The elements of the fundamental set of solutions $y_k(x)$ $[k = 1, 2, \cdots, n]$ can be expressed * in terms of the functions $h_k(x)$ $[k = 1, 2, \cdots, n]$ by the relations

$$y_k(x) = \sum_{m=1}^{n} z_{mj}(x) h_m(x) \qquad [k = 1, 2, \cdots, n], \tag{6.1}$$

in which the functions $z_{ij}(x)$ $[i, j = 1, 2, \cdots, n]$ are periodic of period one and are analytic in Q, and $Z(x) = \text{Det}\,[z_{ij}(x)] \neq 0$. From (6.1), we obtain $y_j(x+i) = \sum_{m=1}^{n} z_{mj}(x+i) h_m(x+i) = \sum_{m=1}^{n} z_{mj}(x) h_m(x+i)$, and hence

$$\begin{aligned} Y(x) &= \text{Det}\,[y_j(x+i)] = \text{Det}\,[\sum_{m=1}^{n} z_{mj}(x) h_m(x+i)] \\ &= \{\text{Det}\,[z_{ij}(x)]\}\,\{\text{Det}\,[h_j(x+i)]\} = Z(x) H(x). \end{aligned}$$

Similarly, we may show that $Y_k(x) = \text{Det}(k)\,[y_j(x+i)] = \sum_{j=1}^{n} H_j(x) Z_{jk}(x)$ in which $Z_{ij}(x)$ is the minor of the element $z_{ij}(x)$ in $Z(x)$. Therefore,

$$w'_k(x) = (-1)^{n+k+1} \sum_{m=0}^{\infty} \frac{b(x+m)}{a_0(x+m)} \, \frac{\sum_{j=1}^{n} H_j(x+m) Z_{jk}(x+m)}{H(x+m) Z(x+m)} \, .$$

But $Z(x)$ and all its minors are periodic of period one, and therefore, we have

$$w'_k(x) = \frac{(-1)^{n+k+1}}{Z(x)} \sum_{m=0}^{\infty} \frac{b(x+m)}{a_0(x+m)} \, \frac{\sum_{j=1}^{n} H_j(x+m) Z_{jk}(x)}{H(x+m)} \, .$$

By hypothesis, the series

---

* See, for example, Batchelder, *loc. cit.*, 9-10.

$$\sum_{m=0}^{\infty} \frac{b(x+m)H_j(x+m)}{a_0(x+m)H(x+m)} \qquad [j=1,2,\cdots,n]$$

converge in Q' and hence converge in Q; therefore the following rearrangement of terms is permissible, and we have

$$w'_k(x) = \frac{(-1)^{n+k+1}}{Z(x)} \sum_{j=1}^{n} Z_{jk}(x) \left[\sum_{m=0}^{\infty} \frac{b(x+m)H_j(x+m)}{a_0(x+m)H(x+m)} \right]$$
$$= \frac{(-1)^{n+k+1}}{Z(x)} \sum_{j=1}^{n} (-1)^{n+j+1} Z_{jk}(x) w_j(x).$$

Therefore,

$$(6.2) \quad y(x) = \sum_{k=1}^{n} w'_k(x) y_k(x)$$
$$= \sum_{k=1}^{n} \left[\frac{(-1)^k}{Z(x)} \sum_{j=1}^{n} (-1)^j Z_{jk}(x) w_j(x) \right] \left[\sum_{i=1}^{n} z_{ik}(x) h_i(x) \right]$$
$$= \frac{1}{Z(x)} \sum_{j=1}^{n} w_j(x) \left[\sum_{i=1}^{n} h_i(x) \sum_{k=1}^{n} (-1)^{j+k} z_{ik}(x) Z_{jk}(x) \right].$$

But since $\sum_{k=1}^{n} (-1)^{j+k} z_{ik}(x) Z_{jk}(x) = Z(x)$ when $i=j$, and equal to zero when $i \neq j$, the bracketed factor of (6.2) is zero, unless $i=j$, and hence (6.2) reduces to

$$y(x) = \frac{1}{Z(x)} \sum_{j=1}^{n} w_j(x) [h_j(x) \sum_{k=1}^{n} (-1)^{j+k} z_{jk}(x) Z_{jk}(x)]$$
$$= \frac{1}{Z(x)} \sum_{j=1}^{n} w_j(x) h_j(x) Z(x) = \sum_{j=1}^{n} w_j(x) h_j(x) = h(x).$$

We have thus shown that $y(x)$ is identical with $h(x)$ in Q, and have thus established the important fact that by using the method of § 5, we are led to the same solution of (1.1) regardless of which fundamental set of solutions of (1.2) are employed.

7. *Asymptotic form.* We shall now examine the asymptotic form of the solution $h(x)$ which was obtained in § 5. Since $h(x) = \sum_{k=1}^{n} w_k(x) h_k(x)$, and since the asymptotic forms of the functions $h_k(x)$ are known, we shall study the forms of the functions $w_k(x)$ $[k=1,2,\cdots,n]$ which are defined by (5.2). The general term of the series

$$(7.1) \qquad w_k(x) = -\sum_{m=0}^{\infty} T_k(x+m) \qquad [k=1,2,\cdots,\gamma]$$

is similar to the general term of (15) of W (p. 215), and the details of the determination of the asymptotic form of (7.1) in the sector

$$(7.2) \qquad -\pi/2 + \epsilon_1 \leq \arg x \leq \pi/2 - \epsilon_1 \qquad (\text{any fixed } \epsilon_1 > 0)$$

are similar to those given in W (pp. 218-220), while the determination of the asymptotic forms of the functions $w_k(x)$ $[k=\gamma+1, \gamma+2, \cdots, n]$ is similar to that given in W (pp. 236-237), and we obtain the result *

$$w_k(x) = (-1)^{n+k+1} \frac{x^\lambda}{(x/e)^{\mu_k x} \rho_k{}^x x^{B_k+r_k}} [\bar{d}_k;\ x],$$

where $\qquad B_k = E - E_k - (\mu_1 + \mu_2 + \cdots + \mu_n),$

$\bar{d}_k = d_k\ [k=1, 2, \cdots, \gamma]$; $\bar{d}_k = (\rho_k d_k)/(\rho_k - 1)\ [k=\gamma+1, \gamma+2, \cdots, n]$.

Since $h_k(x) = (x/e)^{\mu_k x} \rho_k{}^x x^{r_k}[1;\ x]$ in the sector (7.2), we have

$$h(x) = \sum_{k=1}^{n} w_k(x) h_k(x) = \sum_{k=1}^{n} (-1)^{n+k+1} (x^\lambda / x^{B_k}) [\bar{d}_k;\ x].$$

The smallest number of the set B_k $[k=1, 2, \cdots, n]$ is equal to zero, and this value is attained by $B_{\gamma+1} = B_{\gamma+2} = \cdots = B_n = 0$, and hence

$$(7.3) \qquad h(x) = x^\lambda \sum_{k=\gamma+1}^{n} (-1)^{n+k+1} [\rho_k d_k/(\rho_k - 1);\ x].$$

In the case where all of the numbers μ_k are positive, we obtain for the asymptotic form of $h(x)$ the following:

$$(7.4) \qquad h(x) = x^\lambda \sum_{k=\gamma+1}^{n} (-1)^{n+k+1} [d_k;\ x],$$

where γ is defined by $\mu_\gamma \neq \mu_{\gamma+1} = \mu_{\gamma+2} = \cdots = \mu_n$.

By substituting (7.3) or (7.4) into (1.1), we find that the coefficients in the expansion of $\sum_{k=\gamma+1}^{n} (-1)^{n+k+1} [\bar{d}_k;\ x]$ satisfy the same system of equations from which the coefficients $y_0, y_1, y_2, \cdots$ of the formal series solution were uniquely determined, and hence $h(x)$ is asymptotically represented by the formal power series solution $\bar{y}(x)$ in the sector $-\pi/2 < \arg x < \pi/2$. It is also easy to show by direct computation that the first coefficient of the expansions in (7.3) or (7.4) is the same as the first coefficient of the formal power series solution.

Before we examine the asymptotic form of the solution $h(x)$ in regions of the plane other than the sector (7.2), we shall find other forms for this solution. By the lemma of § 6, the series

$$(7.5) \quad w'_k(x) = (-1)^{n+k+1} \sum_{m=0}^{\infty} \frac{b(x+m) G'_k(x+m)}{a_0(x+m) G'(x+m)} = -\sum_{m=0}^{\infty} T'_k(x+m)$$

$$(7.6) \quad w''_k(x) = (-1)^{n+k+1} \sum_{m=0}^{\infty} \frac{b(x+m) H'_k(x+m)}{a_0(x+m) H'(x+m)} = -\sum_{m=0}^{\infty} T''_k(x+m)$$

* We use the notation $[a;\ x]$ to represent the expression $a + b/x + c/x^2 + \cdots + d/x^t + C(x)/x^{t+1}$ where $C(x)$ is uniformly bounded for sufficiently large values of $|x|$ in a region under consideration.

converge in D' which is the finite part of the plane in $|v| > R$, and where

$$G'(x) = \mathrm{Det}\,[g'_j(x+i)], \qquad G'_k(x) = \mathrm{Det}\,(k)[g'_j(x+i)];$$
$$H'(x) = \mathrm{Det}\,[h'_j(x+i)], \qquad H'_k(x) = \mathrm{Det}\,(k)[h'_j(x+i)],$$

and the functions

$$g'(x) = \sum_{k=1}^{n} w'_k(x) g'_k(x); \qquad h'(x) = \sum_{k=1}^{n} w''_k(x) h'_k(x) \tag{7.7}$$

are solutions of (1.1) which are identical with $h(x)$ in the region D'.

Using the second form given for $g'_k(x)$ in (2.3) and proceeding as we did in § 5, we find that

$$G'_k(x) = \frac{(x/e)^{(\mu_1+\mu_2+\cdots+\mu_n)x}(\rho_1\rho_2\cdots\rho_n)^x\, x^{r_1+r_2+\cdots+r_n}}{(x/e)^{\mu_k x}\,\rho_k^x\, x^{r_k}}\, x^{(n-1)\delta} x^{E_k} F'_k(x)$$

in the region D'_r defined by

$$D'_r:\; u \geq \bar{u}, \qquad |v| > R \qquad (\text{any real } \bar{u}),$$

and where $F'_k(x)$ is uniformly bounded in D'_r and approaches zero uniformly as $|v|$ becomes infinite in D'_r. $G'(x)$ is a solution of the first order equation

$$G'(x+1) = (-1)^n \frac{a_n(x+1)}{a_0(x+1)} G'(x) \tag{7.8}$$

and hence is asymptotically represented in the region D'_r by

$$G'(x) = (x/e)^{(\mu_1+\mu_2+\cdots+\mu_n)x}(\rho_1\rho_2\cdots\rho_n)^x\, x^{r_1+r_2+\cdots+r_n} \cdot A x^{E}\,[1+\bar{M}(x)]\,[1+C(x)/v^{t+1}],$$

in which $\bar{M}(x)$ and $C(x)$ are uniformly bounded. This result is obtained by the application of (2.3) to the equation (7.8), remembering that $\delta = 0$ for a first order equation. Therefore

$$\frac{b(x)G'_k(x)}{a_0(x)G'(x)} = \frac{x^{\lambda+(n-1)\delta}}{(x/e)^{\mu_k x}\,\rho_k^x\, x^{r_k+B_k}}\;\frac{b_0}{A a_{0,0}} M'(x), \tag{7.9}$$

in which $M'(x)$ approaches zero uniformly as $|v|$ becomes infinite in D'_r.

In the region D'_l defined by $u \leq \bar{u}$, $|v| > R$, the functions $G'(x)$ and $G'_k(x)$ have the same forms as those given in § 5 for $H(x)$ and $H_k(x)$ respectively.

Let the regions V' and V'' be defined by

$$V':\quad u > u_0, \quad |v| > v_0 \geq R,$$
$$V'':\quad u \leq u_0, \quad |v| > v_0 \geq R,$$

where v_0 is chosen large enough so that

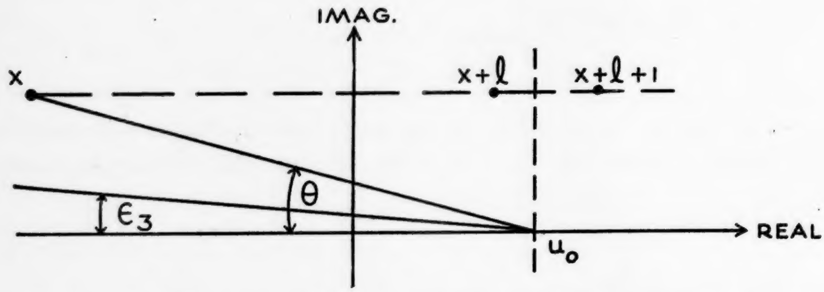
$$(7.10)\quad \begin{array}{lll} \mu_k v + r''_k > 0 & \text{when} \quad v > 0 & [r_k = r'_k + (-1)^{1/2} r''_k], \\ \mu_k v + r''_k < 0 & \text{when} \quad v < 0, & \\ \mid v_0^{\mu_k} \mid > \mid e^{\mu_k}/\rho_k \mid & & [k = 1, 2, \cdots, n], \end{array}$$

and u_0 is chosen so that

$$u_0 > 0; \quad \mu_k u_0 + r'_k + B_k - \lambda - (n-1)\delta > 0 \qquad [k = 1, 2, \cdots, n].$$

In examining the asymptotic form of the series (7.5) in the region V'', we will consider two cases separately; first when x approaches infinity in that part of V'' which is within one of the two sectors

$$(7.11)\qquad \pi/2 + \epsilon_2 \leq \mid \arg (x - u_0) \mid \leq \pi - \epsilon_3,$$



in which ϵ_2 and ϵ_3 are fixed and satisfy $0 < \epsilon_2 < \frac{1}{2} \arcsin (1/4)$, $0 < \epsilon_3 < (1/2) \arctan (1/4)$. Write the series (7.5) in the form

$$(7.12)\qquad w'_k(x) = -\sum_{m=0}^{l} T'_k(x+m) - \sum_{m=l+1}^{\infty} T'_k(x+m),$$

where l is so chosen that the point $(x + l)$ is in V'' while the point $x_1 = x + l + 1$ is in V'. We wish to show that it is possible to restrict ourselves to values of x in V'' and in the sectors (7.11) for which

$$(7.13)\qquad l/\mid x \mid \leq \sigma < 1.$$

The angle ϵ_3 has already been chosen as a fixed positive angle, and since $\cos \epsilon_3 < 1$, there exists a positive number η such that $(1 + \eta) \cos \epsilon_3 \leq \sigma < 1$. The number u_0 is fixed and we can restrict ourselves to those points x for which $\mid x \mid$ is so large that $u_0/\mid x \mid < \eta$. We then see with the aid of the accompanying diagram that

$$\begin{aligned} l &\leq u_0 - u = \mid x - u_0 \mid \cos \theta \\ &\leq (\mid x \mid + u_0) \cos \epsilon_3, \end{aligned}$$

and $l/|x| \leq (1 + u_0/|x|) \cos \epsilon_3 \leq (1 + \eta) \cos \epsilon_3 \leq \sigma < 1$.

Consider the first term of the right hand member of (7.12); it may be written in the form

$$(7.14) \qquad \frac{(-1)^{n+k+1}}{(x/e)^{\mu_k x} \rho_k{}^x x^{c_k}} \sum_{m=0}^{l} \frac{d_k[1 + M_k(x+m)]}{(x/e)^{\mu_k m}(1+m/x)^{\mu_k(x+m)+c_k} \rho_k{}^m}.$$

The first factor is independent of m; we will consider the second factor. Because of (7.13), the quantity $(1 + j/x)^{\mu_k(x+j)+c_k}$ $[j \leq l]$ which appears in the denominator of the $(j+1)$-th term can be developed in a uniformly convergent power series in $1/x$. Hence, we may write *

$$(1 + j/x)^{\mu_k(x+j)+c_k} = e^{\mu_k j}\left[1 + \frac{}{x} + \frac{}{x^2} + \cdots + \frac{}{x^{\mu_k(l-j)}} + \frac{M_{\mu_k(l-j)}(x)}{x^{\mu_k(l-j)}}\right],$$

where $M_{\mu_k(l-j)}(x)$ is uniformly bounded for $|x|$ sufficiently large in V''. Consequently the $(j+1)$-th term of the second factor of (7.14) may be written

$$\frac{1}{\rho_k{}^j x^{\mu_k j}}\left[d_k + \frac{}{x} + \frac{}{x^2} + \cdots \frac{}{x^{\mu_k(l-j)}} + \frac{C_{\mu_k(l-j)}(x)}{x^{\mu_k(l-j)}}\right],$$

and hence (7.14) becomes

$$(7.15) \qquad \frac{(-1)^{n+k+1}}{(x/e)^{\mu_k x} \rho_k{}^x x^{c_k}}\left[d_k + \frac{}{x} + \frac{}{x^2} + \cdots + \frac{}{x^{\mu_k l}} + \frac{C_{\mu_k l}(x)}{x^{\mu_k l}}\right],$$

where $C_{\mu_k l}(x)$ is uniformly bounded for $|x|$ sufficiently large in V''.

Consider now the second term of the right hand member of (7.12); we can write this as

$$(7.16) \qquad -\sum_{m=0}^{\infty} T'_k(x_1 + m)$$
$$= (-1)^{n+k+1} \sum_{m=0}^{\infty} \frac{(x_1+m)^{\lambda+(n-1)\delta}(b_0/Aa_{0,0})M'(x_1+m)}{[(x_1+m)/e]^{\mu_k(x_1+m)} \rho_k{}^{x_1+m} (x_1+m)^{r_k+B_k}},$$

where $x_1 = x + l + 1$, and where we have used (7.9) since the points $x_1, x_1 + 1, x_1 + 2, \cdots$ are in V''. Now,

$$|(x_1 + j)^{\mu_k(x_1+j)+\bar{r}}| = |(x_1)^{\mu_k(x_1+j)+\bar{r}}| \cdot |(1 + j/x_1)|^{\mu_k(u_1+j)+\bar{r}'} e^{-(\mu_k v + \bar{r}'')\arg(1+j/x_1)}$$
$$> |(x_1)^{\mu_k(x_1+j)+\bar{r}}|$$

(where $\bar{r} = \bar{r}' + (-1)^{1/2}\bar{r}'' = r_k + B_k - \lambda - (n-1)^\delta$; $x_1 = u_1 + (-1)^{1/2}v$),

* We shall omit writing coefficients in series in $1/x$ where the coefficients can be determined in terms of given quantities, and their explicit form is of no concern.

since

$$|(1+j/x_1)|>1,\quad \mu_k(u_1+j)+\bar{r}'>\mu_k u_0+\bar{r}'>0,$$

and $\qquad -(\mu_k v+\bar{r}'')\arg(1+j/x_1)=-(\mu_k v+r_k'')\arg(1+j/x_1)>0.$

Therefore, the series (7.16) is dominated by

$$M\sum_{m=0}^{\infty}\frac{1}{|(x_1/e)^{\mu_k(x_1+m)}\rho_k^{x_1+m}x_1^{\bar{r}}|}=\frac{M}{|(x_1/e)^{\mu_k x_1}\rho_k^{x_1}x_1^{\bar{r}}|}\sum_{m=0}^{\infty}\frac{1}{|\rho_k(x_1/e)^{\mu_k}|^m},$$

where M is a constant which is a uniform bound for $b_0M(x_1+m)/Aa_{0,0}$ in V'. Since $\sum_{m=0}^{\infty}1/|\rho_k(x_1/e)^{\mu_k}|^m$ is a geometric series of ratio less than one, by virtue of the third relation of (7.10), its sum is less than some fixed constant $\bar{M}$, and hence (7.16) is less than $M\bar{M}/|(x_1/e)^{\mu_k x_1}\rho_k^{x_1}x_1^{\bar{r}}|$ which in turn is less than $M\bar{M}\,e^{(\mu_k v+r_k'')\arg x}/|\rho_k^{x_1}e^{-\mu_k x_1}|$ since $|x_1|>1$, $\mu_k u_1+\bar{r}'>\mu_k u_0+\bar{r}'>0$, and $(\mu_k v+r'')\arg x>(\mu_k v+r'')\arg x_1$ which is positive. Though we have shown the detailed treatment only for the cases where $\mu_k>0$, the treatment for $\mu_k=0$, $|\rho_k|>1$ would be similar and the only alteration necessary is to replace d_k in (7.15) by $\bar{d}_k$. From (7.7), we have

$$(7.17)\qquad \begin{aligned} g'(x)&=\sum_{k=1}^{n}w'_k(x)g'_k(x)=\sum_{k=1}^{n}g'_k(x)[-\sum_{m=0}^{l}T'_k(x+m)-\sum_{m=l+1}^{\infty}T'_k(x+m)]\\ &=\sum_{k=1}^{n}(-1)^{n+k+1}(x^{\lambda}/x^{B_k})\,[\bar{d}_k;x]-\sum_{k=1}^{n}g'_k(x)\{\sum_{m=l+1}^{\infty}T'_k(x+m)\}.\end{aligned}$$

The second term of the right hand member is dominated by $M'|x|^{\mu_k u}|\rho_k/e^{\mu_k}|^u$, where M' is a constant. Since u becomes negatively infinite as $|x|$ increases in the sectors of (7.11), the above expression vanishes more rapidly than any power of $1/x$, and hence the second term of the right hand member of (7.17) contributes nothing to the asymptotic form of $g'(x)$ and we have as in (7.3) and (7.4), that

$$g'(x)=x^{\lambda}\sum_{k=\gamma+1}^{n}(-1)^{n+k+1}[\bar{d}_k;x],$$

and $g'(x)$ is asymptotically represented by the formal power series solution $\bar{y}(x)$ in the sectors of (7.11).

We shall consider now the asymptotic form of $g'(x)$ in that part of V'' which is in the sectors

$$(7.18)\qquad \pi-2\epsilon_3\le|\arg x|\le\pi.$$

Divide the series for $w'_k(x)$ into two terms as in (7.12), but further divide the first term of the right hand member into the two terms

$$(7.19)\qquad -\sum_{m=0}^{p-1} T'_k(x+m) - \sum_{m=p}^{l} T'_k(x+m),$$

where p is so chosen that $-u/2 < p \leq -u/2+1$, from which we find that $j \leq -u/2+1 \leq |x/2|+1 < |3x/4|$ for $|x|$ sufficiently large, and hence $j/|x| < 3/4$; and consequently the first term of (7.19) can be treated in the same way as we treated the first term of the right hand member of (7.12) and we obtain a similar result.

The second term of (7.19) is dominated by

$$(7.20)\qquad \frac{M}{|\rho_k^x|\cdot|(x+p)^{\mu_k x+r_k+B_k-\lambda}|\cdot|e^{-\mu_k x}|}\sum_{m=p}^{l}\frac{1}{v_0^{\mu_k m}\,|e^{-\mu_k m}\rho_k^m|}.$$

Since, $|x+p|^2=(u+p)^2+v^2 \leq (u/2)^2+v^2=u^2/4+v^2$, and since for all x in the sectors (7.18), we have $|v| \leq |u/4|$, and hence $v^2 \leq u^2/16$, it follows that $|x+p|^2 \leq 5|x|^2/16$; and therefore, $|x+p| < |3x/4|$. Using this and the fact that the sum occurring in (7.20) has a finite number of terms, we find that (7.20) is dominated by

$$M'\,\frac{|(4/3)^{\mu_k x+r_k+B_k-\lambda}|}{|(x/e)^{\mu_k x}\rho_k^x\, x^{r_k+B_k-\lambda}|}.$$

Therefore on multiplying the second term of (7.19) by $g'_k(x)$, we have the product dominated by $\bar{M}(4/3)^{\mu_k u}$, which decreases more rapidly than any power of $1/x$ as $|x|$ becomes infinite in (7.18), and hence contributes nothing to the asymptotic form of $g'(x)$. The treatment of the second term of (7.12) is just as before, and hence $g'(x)$ is asymptotically represented by the formal power series solution $\bar{y}(x)$ also in the sectors (7.18).

We next examine the asymptotic form of $h'(x)$ in the region defined by

$$(7.21)\qquad \pi/2-2\epsilon_1 \leq |\arg x| \leq \pi/2+2\epsilon_2;\quad |v|>R,$$

and find that $h'(x)$ is also represented by $\bar{y}(x)$ in (7.21). Since $g'(x)$ and $h'(x)$ are identical with $h(x)$ and since the regions defined by (7.2), (7.11), (7.18), and (7.21) have overlapping sectors extending to infinity, we have the result that $h(x)$ is asymptotically represented by $\bar{y}(x)$ uniformly for all x outside of a left P-region of sufficiently large radius.

We wish to show next that the solution $h(x)$ has a special kind of uniqueness; we shall show that there is no other analytic solution of (1.1) which is asymptotically represented by the formal power series solution of (1.1) in any domain which has in it at least one horizontal line which extends to infinity to the right. Suppose there existed another solution $\bar{h}(x)$ of (1.1)

which was analytic and asymptotically represented by $\bar{y}(x)$ in a domain as described. This domain would have a region Q in common with the region (7.2) and Q would have at least one horizontal line $\mathcal{L}$ which extends to infinity to the right. The function $f(x)$ defined by $f(x) \equiv \bar{h}(x) - h(x)$ is a solution of the equation (1.2), and hence may be expressed as $f(x) = \sum_{k=1}^{n} \pi_k(x) h_k(x)$ where the functions $\pi_k(x)$ are analytic in Q, and are periodic of period one. The functions $\pi_k(x)$ can be determined from the following system of equations:

$$\sum_{i=1}^{n} \pi_i(x) h_i(x+j) = f(x+j) \qquad [j = 1, 2, \cdots, n],$$

whose coefficients have the determinant $H(x)$. From these equations, we find

$$\pi_j(x) = \frac{1}{H(x)} \begin{vmatrix} h_1(x+1) \cdots h_{j-1}(x+1) \; f(x+1) \; h_{j+1}(x+1) \cdots h_n(x+1) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ h_1(x+n) \cdots h_{j-1}(x+n) \; f(x+n) \; h_{j+1}(x+n) \cdots h_n(x+n) \end{vmatrix},$$

$$[j = 1, 2, \cdots, n].$$

Let $F_{ij}(x)$ denote the minor of the element $f(x+i)$ in the determinant which appears in the numerator of the above fraction. Then

$$\pi_j(x) = \frac{\sum_{i=1}^{n} (-1)^{i+j} f(x+i) \, F_{ij}(x)}{H(x)}.$$

Using (5.3), we readily see that

$$F_{ij}(x) = \frac{(x/e)^{(\mu_1+\mu_2+\cdots+\mu_n)x} (\rho_1\rho_2 \cdots \rho_n)^x \, x^{r_1+r_2+\cdots+r_n}}{(x/e)^{\mu_j x} \rho_j^x \, x^{r_j}} J_{ij}(x),$$

where $J_{ij}(x)$ is the minor of the element in the i-th row and j-th column of the determinant $J(x)$ of (5.4). The expansion of $J_{ij}(x)$ is of the form $x^{\alpha_{ij}} \beta_{ij}(x)$, where α_{ij} is some number less than E of (5.5), and $\beta_{ij}(x)$ is a bounded function for $|x|$ sufficiently large in Q. Using the form of $H(x)$ given in (5.4), we have

$$\pi_j(x) = \sum_{i=1}^{n} (-1)^{i+j} f(x+i) \, \frac{x^{\alpha_{ij}} \beta_{ij}(x)}{(x/e)^{\mu_j x} \rho_j^x \, x^{r_j} A x^E \, [1 + \bar{M}(x)]}.$$

Since $\bar{h}(x)$ and $h(x)$ have the same asymptotic form in Q, the asymptotic form of $f(x)$ in Q is given by $f(x) = x^\lambda [0 + 0 + \cdots + 0 + C_t(x)/x^{t+1}]$, and hence we see that as $|x|$ becomes infinite in Q in such a way that u

becomes positively infinite we have $\lim_{u\to\infty \text{ in } Q} \pi_j(x) = 0$ $[j = 1, 2, \cdots, n]$, and hence in particular $\lim_{|x|\to\infty \text{ on } \bar{L}} \pi_j(x) = 0$. But the functions $\pi_j(x)$ are periodic of period one and hence must be identically zero along $\bar{L}$, and since they are analytic in Q and identically zero along $\bar{L}$ they must be identically zero in Q. Therefore

$$f(x) \equiv \bar{h}(x) - h(x) = \sum_{k=1}^{n} \pi_k(x) h_k(x) \equiv 0,$$

and $\bar{h}(x) \equiv h(x)$.

We collect our results thus far obtained in

THEOREM 1. *When the functions $b(x)$ and $a_k(x)$ $[k = 0, 1, 2, \cdots, n]$ are rational and the numbers μ_k $[k = 1, 2, \cdots, n]$ and the roots of the characteristic equation associated with the horizontal segment of the broken line L satisfy the conditions (5.1), there exists a solution $h(x)$ of the equation (1.1) which is analytic throughout the finite plane except possibly for poles at determinate congruent points within a left P-region. This solution $h(x)$ is asymptotically represented by the formal power series solution of (1.1), uniformly, for all x outside of a left P-region of sufficiently large radius. There is no other analytic solution of (1.1) which is asymptotically represented by the formal power series solution in any domain which has in it at least one horizontal line which extends to infinity to the right.*

The problem which is considered in W may now be regarded as a special case of our problem which is obtained by assuming that all of the numbers μ_k $[k = 1, 2, \cdots, n]$ are equal, and hence our results concerning asymptotic form, which are more extensive than those obtained in W, are also valid for that problem. In pp. 224-227 of W, the asymptotic form of the solution in the left half plane is considered only along rays, whereas we have treated the asymptotic form in the left half plane for any method of approach to infinity. That part of our Theorem 1 pertaining to the uniqueness of $h(x)$ also applies to the first principal solution of W in the case when μ is positive.

8. *The second solution.* By using the functions $g_k(x)$ $[k = 1, 2, \cdots, n]$ as the solutions of (1.2) and the symbolic contour integral solution of equation (3.2), we obtain a second solution of (1.1) given by $g(x) = \sum_{k=1}^{n} \bar{w}_k(x) g_k(x)$, where
where

$$(8.1) \qquad \bar{w}_k(x) = (-1)^{n+k} \int_{\Gamma} \frac{b(z) G_k(z)}{a_0(z) G(z)} \frac{dz}{e^{2\pi(-1)^{1/2}(x-z)} - 1}$$

and $G(x) = \text{Det}\,[g_j(x+i)]$, $G_k(x) = \text{Det}\,(k)\,[g_j(x+i)]$. The functions defined by (8.1) can be shown to exist and be analytic in the finite part of the plane exterior to a right P-region of radius R by methods analogous to those used in W (p. 228). The solution thus obtained can be extended to the right by means of the equation (1.1) itself from which we find that

$$g(x) = \frac{b(x-n)}{a_0(x-n)} - \sum_{k=1}^{n} \frac{a_k(x-n)}{a_0(x-n)}\, g(x-k),$$

which defines $g(x)$ at all finite points within the right P-region which are not singularities of $b(x-n)$ nor $a_k(x-n)$ $[k = 1, 2, \cdots, n]$ nor zeros of $a_0(x-n)$ nor points congruent to all these on the right.

We can also show by methods similar to those used in W (pp. 229-230), that $g(x)$ is asymptotically represented by $\bar{y}(x)$ in the open sector $\pi/2 < \arg x < 3\pi/2$. Our results concerning this second solution are contained in

THEOREM 2. *When the functions $b(x)$ and $a_k(x)$ $[k = 0, 1, 2, \cdots, n]$ are rational and the numbers μ_k $[k = 1, 2, \cdots, n]$ and the roots of the characteristic equation associated with the horizontal segment of the broken line L satisfy the conditions* (5.1), *there exists a solution $g(x)$ of the equation* (1.1) *which is analytic throughout the finite plane except possibly for poles at determinate congruent points within a right P-region. This solution $g(x)$ is asymptotically represented by the formal power series solution of* (1.1) *uniformly in the sector $\pi/2 < \arg x < 3\pi/2$.*

9. If instead of the conditions (5.1) we assume that the following hold

(9.1) $\mu_k = 0,\ |\rho_k| < 1\ [k = 1, 2, \cdots, \gamma]$; $\mu_k < 0\ [k = \gamma+1, \gamma+2, \cdots, n]$,

we find that the rôle of right and left are interchanged. The use of the solutions $g_1(x), g_2(x), \cdots, g_n(x)$; $g'_1(x), g'_2(x), \cdots, g'_n(x)$; $h'_1(x), h'_2(x), \cdots, h'_n(x)$ of equation (1.2) and the symbolic series solution to the left of equation (3.2) are found to yield solutions of (1.1) which we denote by $g(x)$, $g'(x)$, and $h'(x)$, respectively. The solution $g(x)$ is analytic for all finite x exterior to a right P-region of radius R; the solutions $g'(x)$ and $h'(x)$ are shown to exist and are analytic throughout the finite part of the plane in $|v| > R$. These three solutions are identical where they exist and are asymptotically represented by the formal power series solution of (1.1) uniformly for all x outside of a right P-region of sufficiently large radius.

The use of the solutions $h_1(x), h_2(x), \cdots, h_n(x)$ of equation (1.2), and the symbolic contour integral solution of equation (3.2) around a contour extending to infinity to the left as used in W (p. 232) yields another solution

of (1.1) which is analytic throughout the finite part of the plane exterior to a left P-region of radius R, and which is asymptotically represented by the formal power series solution uniformly in the sector $-\pi/2 < \arg x < \pi/2$.

During the investigation of the asymptotic form of these solutions it becomes necessary to prove the following

LEMMA. *Let Q be a region extending to infinity which has the region Q'', extending to infinity, in common with the region Q' into which Q is transformed by the translation $x' = x - n$. If $g(x)$ is an analytic solution of the equation* (1.1) *which is asymptotically represented by $x^q P(1/x)$ in which $P(1/x)$ is a power series in $1/x$, and if $\bar{y}(x) = x^\lambda(y_0 + y_1/x + \cdots)$ is the formal power series solution of* (1.1), *then $q = \lambda$, and $P(1/x) = y_0 + y_1/x + \cdots$; and $g(x)$ is asymptotically represented by $\bar{y}(x)$ in Q.*

Our results are contained in

THEOREM 3. *When the functions $b(x)$ and $a_k(x)$ $[k = 0, 1, 2, \cdots, n]$ are rational and the numbers μ_k $[k = 1, 2, \cdots, n]$ and the roots of the characteristic equation associated with the horizontal segment of the broken line L satisfy the conditions* (9.1), *there exists a solution $g(x)$ of the equation* (1.1) *which is analytic throughout the finite plane except for poles at determinate congruent points within a right P-region. This solution $g(x)$ is asymptotically represented by the formal power series solution of* (1.1), *uniformly, for all x outside of a right P-region of sufficiently large radius. There is no other analytic solution of* (1.1) *which is asymptotically represented by the formal power series solution in any domain which has in it at least one horizontal line extending to infinity to the left.*

There exists a second solution $h(x)$ of the equation (1.1) *which is analytic throughout the finite plane except possibly for poles at determinate congruent points within a left P-region. This solution $h(x)$ is asymptotically represented by the formal power series solution of* (1.1) *uniformly in the sector $-\pi/2 < \arg x < \pi/2$.*

If we assume the following conditions

$$
(9.2)\quad \begin{aligned}
&\mu_k > 0\ [k = 1, 2, \cdots, \alpha];\ \mu_k = 0,\ |\rho_k| > 1\ [k = \alpha + 1, \alpha + 2, \cdots, \beta];\\
&\mu_k = 0,\ |\rho_k| < 1\ [k = \beta + 1, \beta + 2, \cdots, \gamma];\\
&\mu_k < 0\ [k = \gamma + 1, \gamma + 2, \cdots, n],
\end{aligned}
$$

we are able to obtain solutions of (1.1) which are analytic in the finite part of $|v| > R$. This is done by using the symbolic series solution to the right of equation (3.2) for the first β values of k and the symbolic series solution

to the left of equation (3.2) for the values of $k=\beta+1, \beta+2, \cdots, n$. One of the solutions so obtained is asymptotically represented by $\bar{y}(x)$ in the region $0 \leq |\arg x| < \pi/2$, $|v| \geq v_0 > R$, where v_0 is sufficiently large; and the other solution is asymptotically represented by $\bar{y}(x)$ in the region $\pi/2 < |\arg x| \leq \pi$, $|v| \geq v_0 > R$. We are unable by the methods employed thus far and assuming (9.2) to hold to obtain a solution of (1.1) which is analytic throughout the finite plane except possibly for poles in a P-region. In the next section, however, we will employ a method which exhibits the existence of such solutions of (1.1), even with more general assumptions than those in (9.2).

10. *A method applicable to the general case.* If we denote by $I[y(x)]$ the following linear operator on $y(x)$: $I[y(x)] = \sum_{k=0}^{n} a_k(x) y(x+n-k)$, we can write the equations (1.1) and (1.2) in the forms $I[y(x)] = b(x)$ and $I[y(x)] = 0$, respectively. If $y(x)$ is a solution of (1.1), we have $I[y(x+1)] = b(x+1)$, and hence

$$b(x)I[y(x+1)] - b(x+1)I[y(x)] = 0 \tag{10.1}$$

is a linear homogeneous difference equation of the $(n+1)$-th order which is also satisfied by $y(x)$. Let us denote the left hand member of (10.1) by $I'[y(x)]$, so that

$$\begin{aligned} I'[y(x)] &\equiv b(x)I[y(x+1)] - b(x+1)I[y(x)] \\ &= \sum_{k=0}^{n+1} q_k(x) y(x+n+1-k), \end{aligned}$$

where the functions $q_k(x)$ are defined by

$$\begin{aligned} &q_k(x) = b(x)a_k(x+1) - b(x+1)a_{k-1}(x) \quad [k=0,1,2,\cdots,n+1] \\ &a_{-1}(x) = a_{n+1}(x) \equiv 0. \end{aligned} \tag{10.2}$$

We have assumed that the functions $b(x)$ and $a_k(x)$ $[k=0,1,2,\cdots,n]$ are either rational or else of a rational character at infinity; we then see from (10.2) that the functions $q_k(x)$ are of the same nature. We have also restricted ourselves to the study of equation (1.1) under the hypothesis that the equation (1.2) belongs to the type which in A is called class 2a. We shall show that when the equation (1.2) belongs to this type, then also the equation

$$I'[y(x)] = 0 \tag{10.3}$$

belongs to the same type, provided that the characteristic equation associated with the horizontal segment of the broken line L does not have any of its

roots equal to unity. To show that this is true, we must show that the slopes of the segments of the broken line L' associated with the equation (10. 3) are either zero or else positive or negative integers and that each of the several characteristic equations associated with the segments of L' have only simple roots. This is not difficult to show and in fact we shall show that n of the $(n+1)$ pieces * of L' have the same slopes as the n pieces of L, and that the broken line L' has one additional piece whose slope is zero. The characteristic equations associated with those segments of L' which are not horizontal have the same roots as the characteristic equations associated with the corresponding segments of L; if the line L has no horizontal segment, the line L' has a horizontal segment of unit length, and the single root of the characteristic equation associated with this horizontal segment is equal to one; if the line L has a horizontal segment, then the line L' has a horizontal segment which is one unit longer than that of the line L, and the roots of the characteristic equation associated with the horizontal segment of L' include all of the roots of the characteristic equation associated with the horizontal segment of L plus one other root which is equal to one. Hence if the characteristic equation associated with the horizontal segment of L does not have any of its roots equal to unity, then the characteristic equations associated with the line L' have simple roots, and the equation (10. 3) belongs to Adams' class 2a.

Using the expansions of $b(x)$ and $a_k(x)$ as given in (4. 2) and (4. 3), and the relations (10. 2), we find

$$q_k(x) = \frac{x^\lambda}{x^{\mu_{k+1}+\mu_{k+2}+\cdots+\mu_\alpha}}\,(q_{k,0} + q_{k,1}/x + q_{k,2}/x^2 + \cdots) \qquad [k = 0, 1, 2, \cdots, \alpha - 1]$$

$$(10.\,4) \qquad q_k(x) = x^\lambda(q_{k,0} + q_{k,1}/x + q_{k,2}/x^2 + \cdots) \qquad [k = \alpha, \alpha+1, \cdots, \beta, \beta+1],$$

$$q_k(x) = \frac{x^\lambda}{x^{-\mu_{\beta+1}-\mu_{\beta+2}-\cdots-\mu_{k-1}}}\,(q_{k,0} + q_{k,1}/x + q_{k,2}/x^2 + \cdots) \qquad [k = \beta+2, \cdots, n+1],$$

where

$$(10.\,5) \qquad q_{k,0} = b_0(a_{k,0} - a_{k-1,0})\,; \quad a_{-1,0} = a_{n+1,0} = 0 \quad [k = 0, 1, 2, \cdots, n+1],$$

and

$$q_{\alpha,1} = b_1 a_{\alpha,0} + b_0 a_{\alpha,1} - \epsilon_\alpha b_0 a_{\alpha-1,0}\,; \quad \epsilon_\alpha \begin{cases} = 0 \text{ if } \mu_\alpha \neq 1, \\ = 1 \text{ if } \mu_\alpha = 1, \end{cases}$$

---

* We here consider the broken line L made up of n pieces each of whose projections on the i-axis of Figure 1 of A (p. 511) is of unit length.

$$(10.6)\quad q_{k,1} = b_1(a_{k,0} - a_{k-1,0}) + b_0(a_{k,1} - a_{k-1,1}) - \lambda b_0 a_{k-1,0} \qquad [k = \alpha + 1, \alpha + 2, \cdots, \beta],$$

$$q_{\beta+1,1} = - b_1 a_{\beta,0} - b_0 a_{\beta,1} - \lambda b_0 a_{\beta,0} + \epsilon_{\beta+1} b_0 a_{\beta+1,0}; \qquad \epsilon_{\beta+1} \begin{cases} = 0 \text{ if } \mu_{\beta+1} \neq -1, \\ = 1 \text{ if } \mu_{\beta+1} = -1. \end{cases}$$

Comparing the expansions of (10. 4) with those of (4. 2), we see that the line L' associated with the equation (10. 3) has one horizontal piece more than the line L associated with the equation (1. 2), while the other pieces of L' have the same slopes as the n pieces of L. The characteristic equations associated with those segments of L' which are not horizontal have their coefficients equal to b_0 times the corresponding coefficients of the corresponding characteristic equations associated with the non-horizontal segments of L, and hence the roots of these equations are the same. The characteristic equation associated with the horizontal segment of L is

$$(10.7)\qquad \sum_{k=\alpha}^{\beta} a_{k,0}\rho^{\beta-k} = 0.$$

The characteristic equation associated with the horizontal segment of L' is

$$(10.8)\qquad \sum_{k=\alpha}^{\beta+1} q_{k,0}\rho^{\beta+1-k} = 0,$$

and using the relations given in (10. 5), the above equation becomes

$$\sum_{k=\alpha}^{\beta} a_{k,0}\rho^{\beta+1-k} - \sum_{k=\alpha+1}^{\beta+1} a_{k-1,0}\rho^{\beta+1-k} = 0, \quad \text{or} \quad (\rho - 1)\sum_{k=\alpha}^{\beta} a_{k,0}\rho^{\beta-k} = 0,$$

from which we see that the roots of (10. 8) include the roots of (10. 7) plus an additional root which is equal to one. We assume that the equation (10. 7) does not have any of its roots equal to one, and hence (10. 8) has simple roots. We have thus shown that the equation (10. 3) is of the same type as the equation (1. 2).

Every solution of (1. 1) as well as every solution of (1. 2) is also a solution of (10. 3) as can be seen by inspection. We shall show that, conversely, every solution of (10. 3) is either a solution of (1. 2) or else is a periodic function times a solution of (1. 1); and what is more important for our purpose, we shall prove that if S denotes the sets of solutions of (10. 3) which are asymptotically represented by the formal power series solutions of (10. 3) in any region which has in it at least one horizontal line which extends to infinity either to the right or left, then in each fundamental

set of solutions of S there is at least one function which except for a constant multiplier is a solution of (1.1). Let

$$f_1(x), f_2(x), \cdots, f_n(x), f_{n+1}(x) \tag{10.9}$$

be a fundamental set of solutions of (10.3) which are asymptotically represented by the formal power series solutions of (10.3) in the region Q which contains the horizontal line $\bar{L}$ extending to infinity. Then either $I[f_k(x)] = 0$, or $I[f_k(x)] \neq 0$. If the former of these is satisfied, the function $f_k(x)$ is a solution of (1.2); however, not all of the functions of (10.9) can be solutions of (1.2) for the equation (1.2) has only n linearly independent solutions, and hence there is at least one function of the linearly independent set (10.9) for which $I[f_k(x)] \neq 0$. Let $f(x)$ be such a function which is not a solution of (1.2). Then the function $c(x)$ defined by $I[f(x)] = c(x)$ is not identically zero, and exists and is analytic throughout Q. Since $f(x)$ is a solution of (10.3), we have

$$I'[f(x)] = b(x)I[f(x+1)] - b(x+1)I[f(x)] = 0,$$

and hence this imposes the following condition on $c(x)$,

$$b(x)c(x+1) - b(x+1)c(x) = 0,$$

from which we find $c(x+1)/b(x+1) = c(x)/b(x)$ showing that $c(x)/b(x)$ is a periodic function of period one; let us define $\pi(x)$ as this periodic function, then $c(x) = \pi(x)b(x)$, and hence

$$I[f(x)] = \pi(x)b(x). \tag{10.10}$$

We shall show how the function $f(x)$ can be recognized among the functions (10.9). The equation (10.3) has $(n+1)$ formal power series solutions of the form

$$\sigma_k(x) = (x/e)^{\mu_k x} \rho_k^x x^{r_k}(1 + \sigma_{k,1}/x + \sigma_{k,2}/x^2 + \cdots) \quad [k = 1, 2, \cdots, n+1].$$

In one of these formal power series solutions, we have $\mu_k = 0$, $\rho_k = 1$, and hence the formal power series is of the form

$$\sigma(x) = x^r(1 + \sigma_1/x + \sigma_2/x^2 + \cdots). \tag{10.11}$$

The series in (10.11) formally satisfies equation (10.3); substitute (10.11) into (10.3); the resulting equation is

$$\sum_{k=0}^{n+1} q_k(x)(x+n-k)^r [1 + \sigma_1/(x+n-k) + \cdots] = 0.$$

Employ the expansions of $q_k(x)$ as given in (10.4), and remove the common

factor $x^{\lambda+r}$ which occurs in each term. Then on equating to zero the coefficients of x^0 and x^{-1}, we have, respectively,

$$\sum_{k=a}^{\beta+1} q_{k,0} = 0, \tag{10.12}$$

$$\sigma_1 \sum_{k=a}^{\beta+1} q_{k,0} + r \sum_{k=a}^{\beta+1} (n-k) q_{k,0} + \sum_{k=a}^{\beta+1} q_{k,1} + \epsilon_a q_{a-1,0} + \epsilon_{\beta+1} q_{\beta+2,0} = 0, \tag{10.13}$$

where ϵ_a and $\epsilon_{\beta+1}$ are defined in (10.6). Equation (10.12) gives us no new information since we already know that $\sum_{k=a}^{\beta+1} q_{k,0} = 0$, since 1 is a root of the equation (10.8). Hence (10.13) reduces to

$$-r \sum_{k=a}^{\beta+1} k q_{k,0} + \sum_{k=a}^{\beta+1} q_{k,1} + \epsilon_a q_{a-1,0} + \epsilon_{\beta+1} q_{\beta+2,0} = 0. \tag{10.14}$$

Using the relations (10.5) and (10.6), we find that (10.14) reduces to

$$r \sum_{k=a}^{\beta} a_{k,0} - \lambda \sum_{k=a}^{\beta} a_{k,0} = 0. \tag{10.15}$$

By hypothesis, the equation (10.7) does not have any of its roots equal to one, and hence $\sum_{k=a}^{\beta} a_{k,0} \neq 0$, and we find from (10.15) that $r = \lambda$. Therefore the function $f(x)$ is asymptotically represented by

$$\sigma(x) = x^{\lambda}(1 + \sigma_1/x + \sigma_2/x^2 + \cdots) \tag{10.16}$$

in the region Q.

Now returning to (10.10), we have $\sum_{k=0}^{n} a_k(x) f(x+n-k) = \pi(x) b(x)$. Replace the functions $a_k(x)$ and $b(x)$ by their expansions as given in (4.2) and (4.3), and $f(x)$ by its asymptotic form as given by (10.16). The highest degree term on each side of the resulting equation is x^{λ}. Divide both sides of the equation by x^{λ}, and let $|x|$ become infinite along $\bar{L}$, and we find $\lim_{|x| \to \infty \text{ on } \bar{L}} \pi(x) = \text{constant}$. But $\pi(x)$ is a periodic function, and hence it must be identically a constant; let us denote this constant by $1/c$. Hence, $I[f(x)] = b(x)/c$, and $cf(x)$ is a solution of equation (1.1). We therefore have the following

THEOREM 4. *Every solution of the equation* (10.3) *is either a solution of the equation* (1.2) *or else is the product of a periodic function and a*

solution of equation (1.1); *and moreover in each fundamental set of solutions of* (10.3) *which are asymptotically represented by the formal power series solutions of* (10.3) *in any region which has in it at least one horizontal line which extends to infinity either to the right or left, there is one function which except for a constant multiplier is a solution of equation* (1.1).

From the lemma of § 9, we know that this solution $cf(x)$ is asymptotically represented in Q by $\bar{y}(x)$, the formal power series solution of equation (1.1). The presence of this constant multiplier is easily explained. The solutions of the homogeneous equation (10.3) may be multiplied by any arbitrary constant and yet remain solutions; hence the leading constants in the formal series solutions of (10.3) are arbitrary; but the leading constant in the formal series solution of the non-homogeneous equation (1.1) is not arbitrary but uniquely determined. The solution $cf(x)$ of (1.1) is asymptotically represented by $\bar{y}(x)$ and by $c\sigma(x)$ where $\sigma(x)$ is given in (10.16); hence c must be equal to y_0. Therefore, to obtain a solution of (1.1) from among the functions (10.9), it is merely necessary to pick that one which is asymptotically represented by the series (10.16) and multiply it by the constant y_0.

We have thus exhibited the existence of analytic solutions of the non-homogeneous equation (1.1). Among each of the four fundamental sets of solutions which are known as the "principal" solutions and the "intermediate" solutions, there exists a function which except for a constant multiplier is a solution of (1.1). We have the following

THEOREM 5. *When the functions $b(x)$ and $a_k(x)$ $[k=0,1,2,\cdots,n]$ are rational, and the numbers μ_k $[k=1,2,\cdots,n]$ are either positive or negative integers or zero, but none of the roots of the characteristic equation associated with the horizontal segment of L are equal to unity, there exists a solution $h(x)[g(x)]$ of the equation* (1.1) *which is analytic throughout the finite plane except possibly for poles at congruent points within a left [right] P-region. This solution $h(x)[g(x)]$ is asymptotically represented by the formal power series solution of* (1.1) *in the sector $-\pi/2 < \arg x < \pi/2$ $[\pi/2 < \arg x < 3\pi/2]$.*

The methods used in this section for exhibiting analytic solutions of the equation (1.1) also apply in the special cases treated previous to this section, although the results obtained in the previous sections are more extensive than those which are obtained by this method. This method helps to explain why in some of the cases which we considered, we were able to show the uniqueness of one of the solutions which we obtained. When the

numbers μ_k are all either positive or zero [negative or zero], the solution of the equation (1. 1) which we have called $h(x)[g(x)]$ is the principal solution of (10. 3) which is associated with the segment of the broken line L' which is the furthest to the right [left], and Adams has shown in A that these solutions are unique in the sense that there is no other analytic solution of (10. 3) which is asymptotically represented by the formal power series solution in the sector $-\pi/2 < \arg x < \pi/2$ [$\pi/2 < \arg x < 3\pi/2$]. Another point worth noting is that when the conditions (5. 1) [(9. 1)] are satisfied, the solution $h(x)$ [$g(x)$] is the solution which is known as the "determinant limit" solution (See B, pp. 246-255), and from the results in A, we know that $h(x)$ [$g(x)$] is asymptotically represented by the formal power series solution in the sector $-\pi < \arg x < \pi$ [$0 < \arg x < 2\pi$]. The results concerning asymptotic form which are obtained in our Theorems 1 and 3 are however more extensive than the information which has been obtained by treating these solutions as solutions of a homogeneous equation of one higher order.

BROWN UNIVERSITY,
PROVIDENCE. R. I.

CARNEGIE INSTITUTE OF TECHNOLOGY,
PITTSBURGH, PA.

PRIME-POWER ABELIAN GROUPS GENERATED BY A SET OF CONJUGATES UNDER A SPECIAL AUTOMORPHISM.

By H. R. Brahana.

The abelian groups which can be commutator subgroups of groups generated by two operators, S of prime order p and T of order 2, constitute a special class. Among the properties of such groups are the following:

(1) the abelian group H must admit an isomorphism U of order p;
(2) H is generated by the conjugates of one of its operators under powers of U, of which no more than $p-1$ are independent;
(3) H contains no operator of order different from p which is invariant under U.*

As usual the question of the existence of such groups resolves itself into the question of the existence of prime-power groups having the same properties. In the following pages the question is considered in four sections according as the order of H is or is not a power of p and in each case according as H is or is not of type $1, 1, \ldots$ In § 1 attention is called to certain characteristic subgroups of H and certain groups isomorphic with H which are shown to satisfy the above conditions provided H satisfies them. The main results of the paper are necessary and sufficient conditions on H in terms of order and type which are given in (3. 2) and (5. 64). The intricate, though elementary, algebraic computation of § 5 seems not to be avoidable.

1. When H of order q^n is any prime-power group possessing the three properties above there are certain groups depending on H alone which possess the same properties independently of whether or not q is equal to p.

Let H satisfy those conditions, let the operator in the group of isomorphisms I of H whose existence follows from (1) be denoted by U, and let the operator of H given by (2) be denoted by s_1. It is clear that s_1 is an operator of highest order in H, let this order be q^{n_1}.

The operators which are q^m-th powers of operators of H, $m = 0, 1, 2, \cdots, n_1$, constitute a characteristic subgroup of H. Let this subgroup be denoted by H_m, and let the corresponding quotient group of H be denoted by Q_m. It is clear that H_m is transformed into itself by U. Moreover, since

* Groups $\{S, T\}$ whose commutator subgroups are abelian. *Transactions of the American Mathematical Society*, vol. 35 (1933), p. 386.

H_m is in H, it contains no operator of order different from p which is invariant under U, and hence it and U satisfy (3). Also (2) is satisfied by H and U, for some power of s_1 is an operator of highest order in H_m and its conjugates under U generate H_m since conjugates of s_1 under U generate H. Therefore we have

(1.1) *If H is an abelian group of order q^n which satisfies the three conditions of the introduction, then H_m, the group of q^m-th powers in H, is transformed by U in such a way that these conditions are satisfied when H is replaced by H_m.*

The group Q_m is also transformed into itself by U. We shall prove the following theorem.

(1.2) *If H is an abelian group of order q^n which satisfies conditions* (1), (2), *and* (3), *then Q_m, the quotient group of H with respect to its group of q^m-th powers, also satisfies these conditions in which H is replaced by Q_m and U remains the same.*

It is obvious that conditions (1) and (2) hold for Q_m. Let us suppose that condition (3) does not hold. If $q \neq p$, then Q_m will contain an operator q_a of order q which is invariant under U. Corresponding to q_a there will be h_m, where h_m is the order of H_m, operators of H, let one such operator of H be s_a. Then the group $\{H_m, s_a\}$ is invariant under U. If this group is written in co-sets with respect to H_m, each co-set will be invariant under U. Since H_m contains no operator invariant under U, the number h_m is of the form $1 + kp$. H will therefore contain at least q operators invariant under U. This contradicts the assumption that H satisfies (3). Hence Q_m contains no operator except identity invariant under U.

If $q = p$, let us suppose that m_1 is the largest value of m for which Q_{m_1} does not satisfy (3). Let q_a be an operator of order p^2 in Q_{m_1} which is transformed into itself by U. We may assume that H_{m_1} is of type $1, 1, \cdots$ since Q_{m_1+1} does satisfy (3). Then let s_a be an operator of Q_{m_1+1} which corresponds to q_a in Q_{m_1}. It follows that s_a is of order p^3 or p^2, since $q_a^{p^2} = 1$. Moreover, since $U^{-1}s_aU = s_as_\beta$ where s_β is not identity and is in H_{m_1}, it follows that $U^{-1}s_a^pU = s_a^p$, and therefore that s_a is of order p^2.

Now let m_2 be the largest number such that s_a is contained in H_{m_2}. Then the quotient group of H_{m_2} with respect to H_{m_1+1} satisfies the conditions (1), (2), and (3), whereas the quotient group Q' of H_{m_2} with respect to H_{m_1} does not, and Q' contains an operator q'_a corresponding to s_a which is of order p^2 and is invariant under U. Two possibilities arise: the invariant

operator $s_a{}^p$ is either (a) not in a cyclic subgroup of H_{m_2} of order greater than p^2, or (b) it is contained in such a subgroup. In the first case the quotient of Q' with respect to its group of p^2 powers will be of type $2, 2, \cdots 2, 1, 1, \cdots 1$, will be transformed by U according to condition (2), and will contain an operator of order p^2 invariant under U. That this is impossible is shown in § 5 (5.66). In the second case this quotient group will be of the same type, and will contain an operator of order p, corresponding to s_a, invariant under U but not contained in a cyclic group of order p^2. This also is impossible, as a result of (5.66) and (1.1). Hence the theorem is established by the foregoing argument and (5.66) which follows.

2. In this and the following section we shall consider the groups H of order q^n where $q \neq p$, and for the first section shall take H to be of type $1, 1, \cdots$ We shall determine the conditions under which H satisfies (1), (2), and (3).

If α is the exponent to which q belongs, mod p, then the fact that H contains no operator, except identity, invariant under U requires that n be a multiple of α. And since H is generated by a set of conjugates under U, n is not greater than $p-1$. The following theorem is immediately evident.

(2.1) *The abelian group H of order q^a and type $1, 1, \cdots$ admits an isomorphism U of order p such that U and H satisfy conditions* (1), (2), *and* (3).

For, the group of isomorphisms I of H contains operators of order p. If we designate one such by U, the conjugates of an operator s_1 of H under U generate a subgroup of H which is transformed into itself by U. The order of this group cannot be less than q^a since no group of order q^k, $k < \alpha$, admits an automorphism of order p.

These considerations prove the much stronger theorem:

(2.2) *Any operator of order p in the group of isomorphisms of the group H of order q^a and type $1, 1, \cdots$ transforms any operator of H successively into a set of operators which generate H.*

From a different point of view we get immediately:

(2.3) *The abelian group H of order q^{p-1} and type $1, 1, \cdots$ admits an isomorphism U of order p such that H and U satisfy conditions* (2) *and* (3).

The following isomorphism obviously is the required one. Let a set of generators of H be $s_1, s_2, \cdots, s_{p-1}$. Then let

(2.4) $U^{-1}s_iU = s_{i+1}$, $i = 1, 2, \cdots, p-2$, $U^{-1}s_{p-1}U = s_1{}^{-1}s_2{}^{-1} \cdots s_{p-1}^{-1}$.

The isomorphism U exists. It is of order p since

$$U^{-p}s_1U^p = U^{-1} \cdot s_1^{-1}s_2^{-1} \cdots s_{p-1}^{-1} \cdot U = s_1.$$

The conjugates of s_1 under U are a set of generators.* If $s = s_1^{x_1}s_2^{x_2} \cdots s_{p-1}^{x_{p-1}}$ is invariant under U, we have

$$U^{-1}sU = s_2^{x_1}s_3^{x_2} \cdots s_{p-1}^{x_{p-2}} \cdot s_1^{-x_{p-1}}s_2^{-x_{p-1}} \cdots s_{p-1}^{-x_{p-1}} = s.$$

From this we get the following set of homogeneous congruences:

$$(2.5)\quad \begin{array}{l} x_1 \qquad\qquad\qquad\qquad + x_{p-1} \equiv 0, \\ -x_1 + x_2 \qquad\qquad\qquad + x_{p-1} \equiv 0, \\ \quad -x_2 + x_3 \qquad\qquad + x_{p-1} \equiv 0, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \qquad\qquad -x_{p-3} + x_{p-2} + x_{p-1} \equiv 0, \\ \qquad\qquad\qquad -x_{p-2} + 2x_{p-1} \equiv 0, \bmod q. \end{array}$$

The rank of the matrix of coefficients of (2. 5) is $p-1$ and hence the system has no solution except $0, 0, \cdots 0$, unless $q = p + 1$, in which case $p = 2$ and $q = 3$. H and U satisfy (3) when $q \neq 3$. When $q = 3$, U may be taken to be the operator of order 2 which transforms every operator of H into its inverse. Therefore (2. 3) is always true.

We may observe for future reference that the isomorphism U defined by (2. 4) exists, is of order p, and satisfies (2) whenever H has $p-1$ independent generators of the same order, regardless of whether or not $q = p$. The operators left invariant by U are in general determined by the set of congruences (2. 5) taken modulo q^{n_1}; when $q^{n_1} = p$, the rank of the matrix of coefficients is $p-2$, and there is one and only one solution of the homogeneous system.

The theorems (2. 1) and (2. 3) are special cases of a theorem which we shall now undertake to prove.

Let H be of order q^{ka} and type $1, 1, \cdots$ The group of isomorphisms I of H is of order

$$q^{ka(ka-1)/2}(q^{ka}-1)(q^{ka-1}-1)\cdots(q-1).$$

Of these factors just k are divisible by p, viz:

$$(q^{ka}-1),\ (q^{(k-1)a}-1), \cdots, (q^a-1).$$

Therefore the Sylow subgroup of I corresponding to the prime p is of order p^{km} where p^m is the highest power of p which divides $q^a - 1$.† We

---

* Here and throughout we understand the *set of conjugates of* s_1 to include s_1. When later we use the expression *the m-th conjugate of* s_1 *under* U we shall understand s_1 to be the first conjugate, the m-th conjugate is $U^{-(m-1)}s_1U^{m-1}$.

† For example, if $q = 17$ and $p = 3$, then $a = 2$ and $m = 2$.

may separate the $k\alpha$ generators of H into k sets of α generators each. Let the group generated by $s_1, s_2, \cdots, s_\alpha$ be denoted by H_1, let $\{s_{\alpha+1}, s_{\alpha+2}, \cdots, s_{2\alpha}\}$ be H_2, etc. Then there exist subgroups $U_1, U_2, \cdots, U_k$ of order p^m in the group of isomorphisms I of H, each of which transforms the operators of one of the H_i's and leaves fixed all the operators of each of the others. These U_i's are cyclic, since each is a Sylow subgroup of the group of isomorphisms I_i of the group H_i and I_i contains a cyclic group of order $q^\alpha - 1$.* Then u_iu_j, where u_i is any operator of U_i and u_j is any operator of U_j, performs the same transformation on H as u_ju_i, since they transform the generators of H identically. Therefore,

(2.6) *The Sylow subgroup corresponding to the prime p in the group of isomorphisms of the abelian group of order $q^{k\alpha}$ and type $1, 1, \cdots$, where q belongs to α and $k\alpha \leqq p-1$, is abelian and of type $m, m, \cdots$, where p^m is the highest power of p contained in $q^\alpha - 1$.*

In general I has more than one Sylow subgroup of order p^{km}, but since they are all conjugate, if I contains an operator of order p which transforms an operator of H into a set of generators, then it will contain an operator $U = u_1u_2 \cdots u_k$ where u_i is in U_i and is of order p, which transforms some operator of H into a set of generators.

Any operator U_1 of order p in the group of isomorphisms I_1 of H_1 transforms s_1 into a set of operators which generate H_1. It is then possible to express the $\alpha + 1$-st conjugate in terms of the α preceding ones. The generators of H_1 may then be so chosen that

$$(2.7)\qquad \begin{aligned} U_1^{-1}s_iU_1 &= s_{i+1}, \qquad (i = 1, 2, \cdots, \alpha - 1), \\ U_1^{-1}s_\alpha U_1 &= s_1^{a_1}s_2^{a_2}s_3^{a_3} \cdots s_\alpha^{a_\alpha}. \end{aligned}$$

If U'_1 is any other operator of order p in I_1 we may choose a notation so that

$$\begin{aligned} U'^{-1}_1 s'_i U'_1 &= s'_{i+1} \qquad (i = 1, 2, \cdots, \alpha - 1), \\ U'^{-1}_1 s'_\alpha U'_1 &= s'^{b_1}_1 s'^{b_2}_2 \cdots s'^{b_\alpha}_\alpha. \end{aligned}$$

If the operators U_1 and U'_1 are conjugate in I_1, then the exponents $b_1, b_2, \cdots, b_\alpha$ are obviously the same as the exponents $a_1, a_2, \cdots, a_\alpha$. The converse is also true. For, the isomorphism of H_1 which transforms s_i into s'_i $(i = 1, 2, \cdots, \alpha)$, transforms U_1 into U'_1. Hence

(2.8) *A necessary and sufficient condition that two operators of order p in*

* Moore, "Concerning Jordan's linear groups," *Bulletin of the American Mathematical Society*, Ser. 2, Vol. 2 (1895), p. 33.

the group of isomorphisms of the abelian group of order q^α and type 1, 1, $\cdots$ *be conjugate, is that the first* $\alpha + 1$ *conjugates of an operator* s_1 *of the abelian group under one of those operators be connected by the same relation as the first* $\alpha + 1$ *conjugates of some* s'_1 *under the other.*

A simple isomorphism may be established between H_1 and H_i and by its means U_1 will determine an operator U_i of the group of isomorphisms I_i of H_i. Then as a result of (2. 6) it follows that any operator of order p of I is conjugate to $U = U_1^{a_1}U_2^{a_2}\cdots U_k^{a_k}$, for a proper choice of $a_1, a_2, \cdots, a_k$. Since U_1 is any operator of order p in I_1 we may suppose that $a_1 = 1$. Let us suppose that there exists some operator s in H whose conjugates under U generate H. Then since every operator of H is expressible in only one way in terms of generators of $H_1, H_2, \cdots, H_k$, and since U leaves H_i invariant, it follows that s must be the product of k operators, one from each H_i and none of them identity. Since the conjugates of any operator of H_i under U_i generate H_i, we may choose the generators of H_i so that $s = s_1 s_{\alpha+1}\cdots s_{(k-1)\alpha+1}$, and so that U_i is determined by a set of relations exactly like (2. 6) except that the subscripts on the s's run through α integers beginning with $(i-1)\alpha+1$, and the exponents in the second line depend on i. Let the i-th set of exponents be $a_{i1}, a_{i2}, \cdots, a_{i\alpha}$; also let r_{il} and r_{jl} denote the transforms of $s_{(i-1)\alpha+1}$ and $s_{(j-1)\alpha+1}$ respectively by U^l. If the ordered set $a_{i1}, a_{i2}, \cdots, a_{i\alpha}$ is the same as the set $a_{j1}, a_{j2}, \cdots, a_{j\alpha}$, $i \neq j$, then any combination of the conjugates of s under U will contain r_{il} and r_{jl} to the same power. The group generated by conjugates of s cannot contain either H_i or H_j, and therefore cannot be H. Since $U_i^{a_i}$ is conjugate to $U_1^{a_i}$ in I, we have the following theorem:

(2. 9) *If the group of isomorphisms of the group of order* $q^{k\alpha}$ *and type* 1, 1, $\cdots$ *contains an operator of order p which transforms one of its operators into a set of generators, then the operators of order p in the group of isomorphisms of the group of order* q^α *and type* 1, 1, $\cdots$ *belong to at least k conjugate sets.*

Conversely, if the operators of order p of U_1 belong to k conjugate sets, where $k \leq (p-1)/\alpha$, then the operator U exists and transforms s into a set of generators of the group of order $q^{k\alpha}$ and type 1, 1, $\cdots$ Hence,

(2. 10) *If the operators of order p of the group of isomorphisms of the group of order* q^α *and type* 1, 1, $\cdots$ *belong to k conjugate sets, where* $k \leq (p-1)/\alpha$, *then the group of isomorphisms of the group H of order* $q^{k\alpha}$

and type $1, 1, \cdots$ *contains an operator of order* p *which transforms an operator of* H *into a set of generators.*

From (2.3) and (2.9) we have

(2.11) *The operators of order* p *in the group of isomorphisms of the group of order* q^a *and type* $1, 1, \cdots$ *belong to at least* $(p-1)/\alpha$ *conjugate sets.*

From (2.10) and (2.11) we have

(2.12) *There exists an operator* U *of order* p *in the group of isomorphisms of the group* H *of order* $q^{k\alpha}$, $k \leq (p-1)/\alpha$, *and type* $1, 1, \cdots$ *which transforms an operator of* H *into a set of generators.*

The group H and the operator U of the last theorem satisfy condition (3). For suppose $U = U_1^{a_1} U_2^{a_2} \cdots U_k^{a_k}$ and s is designated as in the proof of (2.9). Then we may adjoin to H certain other groups $H_{k+1}, H_{k+2}, \cdots, H_{(p-1)/\alpha}$ each of order q^a and type $1, 1, \cdots$ to obtain a group of order q^{p-1} and type $1, 1, \cdots$; and we may multiply U by operators $U_{k+1}^{a_{k+1}}, U_{k+2}^{a_{k+2}}, \cdots$, each of order p, to obtain an operator V of order p which transforms an operator of the new group H' into a set of generators. For a proper choice of the generators of H', V will take the form (2.4), for V may be written

$$V^{-1} s_i V = s_{i+1}, \qquad (i = 1, 2, \cdots, p-2),$$
$$V^{-1} s_{p-1} V = s_1^{a_1} s_2^{a_2} \cdots s_{p-1}^{a_{p-1}}.$$

The condition that V be of order p requires that $a_1 = a_2 = \cdots a_{p-1} = -1$. Since the congruences (2.5) have no solution it follows that H' contains no operator, except identity, invariant under V. Since H is transformed by V in the same manner as by U it follows that H contains no operator, except identity, invariant under U. We have therefore the promised theorem which includes (2.1) and (2.3).

(2.13) *The group of isomorphisms of the group of order* $q^{k\alpha}$, $k \leq (p-1)/\alpha$, *and type* $1, 1, \cdots$ *contains operators* U *of order* p *which satisfy with* H (1) *and* (2), *and every such* U *with* H *satisfies* (3).

3. In the present section we suppose H not to be of type $1, 1, \cdots$. As a result of (1.1), (1.2), and the fact that the group of isomorphisms of the group of order q^n and type $1, 1, \cdots$ contains an operator of order p satisfying the conditions of the introduction only if n is a multiple of α, we may suppose that H is of order $q^{\alpha(k_1 m_1 + k_2 m_2 + \cdots + k_j m_j)}$ with $k_i \alpha$ generators of order q^{m_i}. And since H satisfies (2) we may suppose that

$$k_1 + k_2 + \cdots + k_j = k \leq (p-1)/\alpha.$$

The independent generators of H may be grouped, as in § 2, in k sets of α each, where those in one set are all of the same order. Let us denote the group generated by those of the i-th set by H_i. Let the characteristic subgroup of order q^a and type $1, 1, \cdots$ of H_i be denoted by H'_i. Then an operator of order p in the group of isomorphisms I'_i of H'_i determines an operator of order p in the group of isomorphisms I_i of H_i. To see this we may set up a correspondence between generators s_i of H_i and s'_i of H'_i in which s'_i is in the cyclic group generated by s_i. An operator of order p of I'_i thereby determines one or more operators of I_i whose orders are all multiples of p. At least one of those orders is a power of p; * let such an operator be U_i. Then $V_i = U_i^p$ is of order a power of p, and it leaves every operator of H'_i fixed. Let s_a be an operator of H_i such that $V_i^{-1} s_a V_i = s_a s_\beta$ where $s_\beta \neq 1$, and such that $V_i^{-1} s_a^q V_i = s_a^q$. Then s_β is in H'_i. Consequently $V_i^{-p} s_a V^p = s_a s_\beta^p$. Now the order of V_i being a power of p requires s_β to be identity and hence every operator of H_i is invariant under V_i. Therefore U_i is of order p.

It follows from (2.2) that U_i transforms any operator of highest order of H_i into a set of generators of H_i. It is obvious that two non-conjugate operators of order p of I'_i determine two non-conjugate operators of I_i. Moreover, since H'_i and H'_j are identical we may drop the subscripts.

Now I' contains $(p-1)/\alpha$ operators of order p belonging to distinct conjugate sets. Let us denote k of these operators by $U'_1, U'_2, \cdots, U'_k$. Then let the operator in I_i determined by U'_i be U_i. A choice of generators of H_i may be made so that U_i is defined by a set of relations the same as (2.7) except that the subscripts on the s's run through α consecutive integers beginning with $(i-1)\alpha + 1$, and as a result of (2.8) the ordered set of exponents in the second line is not the same as that for $U_j, j \neq i$. Let the i-th set of exponents be $a_{i1}, a_{i2}, \cdots, a_{ia}$. The operator $U = U_1 U_2 \cdots U_k$ is of order p and is in I.

The $(\alpha + 1)$-st conjugate of $s = s_1 s_{a+1} \cdots s_{(k-1)a+1}$ under U is

$$(3.1) \qquad s_1^{a_{11}} s_2^{a_{12}} \cdots s_a^{a_{1a}} s_{a+1}^{a_{21}} s_{a+2}^{a_{22}} \cdots s_{2a}^{a_{2a}} \cdots s_{(k-1)a+1}^{a_{k1}} s_{(k-1)a+2}^{a_{k2}} \cdots s_{ka}^{a_{ka}}.$$

By a proper combination of the first $\alpha + 1$ conjugates of s we may obtain an operator of the form (3.1) where $a_{11} = a_{12} = \cdots = a_{1a} = 0$, and no other set $a_{i1}, a_{i2}, \cdots a_{ia}$ consists solely of zeros or multiples of q, for then U'_1 and U'_i would be conjugate, (2.8). This is an operator of the same type as s, that is, it is the product of operators of highest order, one from each $H_i, i = 2, 3, \cdots, k$. Let the operator which belongs to H_i be denoted by

* Miller, Blichfeldt, and Dickson, *Finite Groups*, p. 67.

$t_{(i-1)a+1}$ and let $t = t_{a+1}t_{2a+1} \cdots t_{(k-1)a+1}$. Then by (2. 2) the conjugates of $t_{(i-1)a+1}$ under U_i generate H_i. The generators of H_i may be selected anew so that U_i is defined by an expression of the form (2. 7) in which the ordered set of exponents in the second line is distinct from that for $U_j, j \neq i$. A proper combination of the first $\alpha + 1$ conjugates of t under U gives an operator of the form (3. 1) where $a_{ij} = 0$, for $i = 1, 2$ and every j. This process may then be continued and after $k - 1$ steps we arrive at an operator of highest order in H_k. As a result of (2. 2) it follows that H_k is in the group generated by the conjugates of s. Since in the above argument no use was made of any special properties of H_k it follows that $H_i, i = 1, 2, \cdots, k$, is in the group generated by the conjugates of s, and therefore that this group is H. Moreover, since the operator U was obtained precisely as in the manner preceding (2. 9) it follows that the subgroup of H of order q^{ka} and type $1, 1, \cdots$ is generated by the conjugates under U of one of its operators and hence by (2. 13) contains no operator invariant under U. Therefore, H contains no operator invariant under U. We have proved the theorem:

(3. 2) *Necessary and sufficient conditions that an abelian group of order q^n admit an automorphism U of order p such that H and U satisfy conditions* (1), (2), *and* (3) *are that $n = \alpha(k_1m_1 + k_2m_2 + \cdots + k_jm_j)$ and that H have $k_i\alpha$ independent generators of order q^{m_i}, where the m's and k's are subject only to the relation $k_1 + k_2 + \cdots + k_j \leq (p - 1)/\alpha$.*

The operators $U'_1, U'_2, \cdots, U'_k$ have been shown to exist; it is perhaps worth while to notice a simple method to find them. They are all in a cyclic group of I'. If there exists any operator of this group not conjugate to U'_1, then U'^{β}_1 must be such, where β is any primitive root, mod p.

4. A group H of order p^k and type $1, 1, \cdots$ has operators of order p in its group of isomorphisms. When $k = p - 1$ one these operators of order p transforms some operator s_1 of H into a set of generators of H (cf. (2. 4) and ff.), and on the other hand if the group of isomorphisms I contains an operator U of order p which transforms s_1 into a set of generators, then $k \leq p - 1$.

We shall suppose that $k \leq p - 1$ and that U exists. Then H contains a subgroup of order p whose operators are invariant under U. We shall show that H contains but one such subgroup. According to the method of § 2 we may select the generators of H so that U is defined as follows:

$$(4. 1') \qquad \begin{aligned} U^{-1}s_iU &= s_{i+1}, \qquad (i = 1, 2, \cdots, k - 1), \\ U^{-1}s_kU &= s_1^{a_1}s_2^{a_2} \cdots s_k^{a_k}. \end{aligned}$$

It will be convenient to designate (4. 1′) by means of the matrix of the exponents on the right-hand side and write

$$(4.1)\qquad U = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ a_1 & a_2 & a_3 & a_4 & \cdots & a_{k-1} & a_k \end{pmatrix}.$$

If $s = s_1^{x_1}s_2^{x_2}\cdots s_k^{x_k}$ is an operator left invariant by U, we have

$$(4.2)\qquad U^{-1}sU = s_2^{x_1}s_3^{x_2}\cdots s_k^{x_{k-1}}\cdot s_1^{a_1x_k}s_2^{a_2x_k}\cdots s_k^{a_kx_k} = s.$$

From this we get the set of homogeneous congruences

$$(4.3)\qquad \begin{array}{l} x_1 \qquad\qquad\qquad\qquad\qquad\qquad\qquad - a_1x_k \equiv 0 \\ -x_1 + x_2 \qquad\qquad\qquad\qquad\qquad\quad - a_2x_k \equiv 0 \\ \qquad - x_2 + x_3 \qquad\qquad\qquad\qquad\quad - a_3x_k \equiv 0 \\ \qquad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot \\ \qquad\qquad\qquad\qquad - x_{k-2} + x_{k-1} - a_{k-1}x_k \equiv 0 \\ \qquad\qquad\qquad\qquad\quad - x_{k-1} + (1 - a_k)x_k \equiv 0, \bmod p. \end{array}$$

In order that (4. 3) have a solution it is necessary and sufficient that the rank of the matrix

$$(4.4)\qquad \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & -a_1 \\ -1 & 1 & 0 & \cdots & 0 & 0 & -a_2 \\ 0 & -1 & 1 & \cdots & 0 & 0 & -a_3 \\ 0 & 0 & -1 & \cdots & 0 & 0 & -a_4 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & -1 & 1 & -a_{k-1} \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 - a_k \end{pmatrix}$$

be less than k. This requires that

$$(4.5)\qquad a_1 + a_2 + \cdots + a_k \equiv 1, \bmod p.$$

The rank of (4. 4) is actually $k - 1$ if (4. 5) is satisfied and hence the system (4. 3) has but one solution $(x_1, x_2, \cdots, x_k)$ different from $(0, 0, \cdots, 0)$.

The above argument is independent of the order of U and we have therefore the following theorem:

(4. 6) *If U is an operator in the group of isomorphisms of H which trans-*

forms an operator of H into a set of generators then H contains at most one subgroup of order p whose operators are invariant under U, and contains exactly one such subgroup when the order of U is p.

Moreover, any set of x's for which $x_k \neq 0$ determines by means of the congruences (4.3) a set of numbers $a_1, a_2, \cdots, a_k$ which determines an operator of the form (4.1). Consequently, if I contains any operator U of order p which transforms an operator of H into a set of generators, then U determines a subgroup of order p in H, and this subgroup is transformed into itself by just one operator of order p of the form (4.1) when the exponent x_k in the expression for s is not zero. Let the subgroup corresponding to U be denoted by C_u.

What we want for our immediate purpose is the fact that for a given choice of generators of H there is just one operator of order p and of the form (4.1) in I. Let U and U' be two operators of the type in question. Consider two sets of generators $s_1, s_2, \cdots, s_k$ and $s'_1, s'_2, \cdots, s'_k$ chosen so that U and U' both take the form (4.1). The two sets of numbers $a_1, a_2, \cdots, a_k$ and $a'_1, a'_2, \cdots, a'_k$ are conceivably different, in which case it follows from the proof of (4.6) that U and U' leave invariant respectively two subgroups which are not conjugate under the isomorphism which transforms s_i into s'_i. Let us then transform C'_u into C_u, and let U'' be the operator of I into which U' is thereby transformed. Then a set of generators $s_1'', s_2'', \cdots, s_k''$ may be selected so that U'' takes the form (4.1). The isomorphism which transforms s_i'' into $s_i, i = 1, 2, \cdots, k$, transforms U'' into U, since both are of order p and leave C_u invariant. The numbers $a_1'', a_2'', \cdots, a_k''$ are therefore the same as $a_1, a_2, \cdots, a_k$. The a'''s are the same as the a''s, for s_1'' may be selected as the conjugate of s'_1 under the operator which transforms C'_u into C_u and $s_2'', s_3'', \cdots, s_k''$, as successive transforms of s_1'' by U''. The relation expressing the $k+1$-st conjugate of s_1'' in terms of the preceding k conjugates will be the same as that for the corresponding conjugates of s'_1, under U'. Therefore we have the theorem:

(4.7) *When H is of order p^k and type $1, 1, \cdots$ every operator of order p in I which transforms an operator of H into a set of generators is conjugate to* (4.1) *in which the a_i's depend only on k and p.*

As has been observed, the number k is not greater than $p-1$, and it was shown in § 2 that when $k = p-1$ the set of numbers $a_1, a_2, \cdots, a_k$ is $-1, -1, \cdots, -1$. These a's substituted in (4.3) determine the set of x's giving the subgroup of H composed of operators invariant under U to be $1, 2, 3, \cdots, p-1$. Hence, $s = s_1 s_2^2 s_3^3 \cdots s_{p-1}^{p-1}$. The quotient group of

H with respect to the group generated by s is of order p^{p-2} and type $1, 1, \cdots$ and it is transformed by U according to an operator of order p which transforms one of its operators into a set of generators. The $(p-1)$-st conjugate of this operator is expressible in terms of the preceding conjugates and the expression for it is obtained immediately by setting s equal to identity. We thus have $s_1\ s_2{}^2\ s_3{}^3 \cdots s_{p-1}^{p-1} = 1$, and $s_{p-1} = s_1\ s_2{}^2 \cdots s_{p-2}^{p-2}$. Therefore if $k = p-2$, the set of a's in (4.1) is $1, 2, \cdots, p-2$. This set of a's used in (4.3) determines the s for $k = p-2$, which in turn determines the $(p-2)$-nd conjugate of s_1 in terms of the preceding conjugates and hence gives the set of a's in (4.1) for $k = p-3$.

The set of a's for $k = p-2$ may be written as $-(p-1), -(p-2), \cdots, -3, -2$, or $-\binom{p-1}{1}, -\binom{p-2}{1}, \cdots, -\binom{3}{1}, -\binom{2}{1}$, these last numbers being the binomial coefficients. We shall show by induction that when $k = p-1-r$ the set of a's in (4.1) is

$$-\binom{p-1}{r}, -\binom{p-2}{r}, \cdots, -\binom{r+1}{r}. \tag{4.8}$$

We assume that (4.8) is correct for a given k and show that for the next smaller value of k we get (4.8) with r replaced by $r+1$. These a's in (4.8) when substituted in (4.3) determine the x's for the group invariant under U. If we solve for them successively in terms of x_k beginning with the last equation we have

$$\begin{aligned}
x_{k-1} &= (1-a_k)x_k = \left[\binom{r}{r} + \binom{r+1}{r}\right]x_k = \binom{r+2}{r+1}x_k,\\
x_{k-2} &= x_{k-1} - a_{k-1}x_k = \left[\binom{r+2}{r+1} + \binom{r+3}{r}\right]x_k = \binom{r+3}{r+1}x_k,\\
&\cdots\cdots\cdots\cdots\\
x_{k-j} &= \binom{r+j+1}{r+1}x_k.\\
&\cdots\cdots\cdots\cdots\\
x_1 &= \binom{p-1}{r+1}x_k, \quad \operatorname{mod} p.
\end{aligned} \tag{4.9}$$

Hence the invariant subgroup is generated by

$$s = \left[s_1{}^{\binom{p-1}{r+1}} s_2{}^{\binom{p-2}{r+1}} \cdots s_{k-1}{}^{\binom{r+2}{r+1}} s_k\right]^{x_k}.$$

For the next quotient group we set s equal to identity. This gives s_k in terms of the preceding ones and determines the a's of (4.1) to be (4.8) with r replaced by $r+1$.

We state explicitly in the following theorem the principal result of this section.

(4.10) *Every abelian group of order p^k, $k \leq p-1$, and type $1, 1, \cdots$ satisfies the conditions of the introduction.*

5. We now consider groups H of order p^n and not of type $1, 1, \cdots$ and we take first those of order p^{km} and type $m, m, \cdots$. When $k = p - 1$ the argument following (2.3) applies, even though the generators of H are of order p^m, and hence I contains an operator U of order p which transforms an operator of H into a set of generators. This operator is given by (4.1′) and by (4.1), where the a's are $-1, -1, \cdots, -1$, mod p^m instead of mod p. From the argument preceding (4.6) it follows that H contains no operator of order greater than p invariant under U. U transforms the group of order p^k and type $1, 1, \cdots$ according to an operator of order p and hence by (4.6) the operators of H invariant under U constitute a subgroup of order p.

This invariant subgroup is determined by the set of congruences (4.3) taken with the modulus p^m, $a_i = -1$, $i = 1, 2, \cdots, k(-p-1)$. The sum of the k congruences is $-px_{p-1} \equiv 0$, mod p^m. Hence, $x_{p-1} \equiv 0$, mod p^{m-1}. The subgroup is generated by

$$(5.1) \qquad (s_1 s_2^2 s_3^3 \cdots s_{p-1}^{-1})^{p^{m-1}}.$$

The quotient group of H with respect to the group invariant under U is of order p^{km-1} and type $m, m, \cdots, m, m-1$, and is transformed by U according to an operator of order p which transforms one of its operators into a set of generators. We may take the generators of this quotient group H_1 to be $s'_1, s'_2, \cdots, s'_{p-1}$ all of order p^m but not all independent. We may suppose them to be successive conjugates under U in which case the first $p-2$ of them will be independent and the expression for $(s'_{p-1})^{p^{m-1}}$ is given by setting (5.1) equal to identity after replacing s_i by s'_i.

A set of independent generators is obtained by taking $s'_1, s'_2, \cdots, s'_{p-2}$ and r_{p-1}, where r_{p-1} is an operator of order p^{m-1} determined by the relation $s'_{p-1} = s'_1 s'^2_2 \cdots s'^{p-2}_{p-2} r_{p-1}$. The transformation U may now be expressed in terms of the independent generators of H_1. It will be noted that the $(p-2)$-rowed square matrix in the upper left hand corner of the matrix for U will be the same as the matrix (4.1) for $k = p-2$, since the group of order p^{k-1} and type $1, 1, \cdots$ composed of operators which are p^{m-1}-th powers in H_1 must be transformed in that manner. This matrix is

$$(5.2) \qquad \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 1 & 2 & 3 & \cdots & p-2 & 1 \\ -p & -2p & -3p & \cdots & -(p-3)p & -(p-1) \end{pmatrix}.$$

The last row gives the transform of r_{p-1} by U and is determined by the definition of r_{p-1} in terms of the s''s.

This group H_1 of order p^{km-1} is transformed by U according to (5.2). The characteristic subgroup of order p^{k-1} and type $1, 1, \cdots$ described above is then transformed by U according to an operator of order p of the type considered in § 4. H_1 then contains a subgroup of order p composed of operators invariant under U, and the quotient group H_2 of H_1 with respect to this subgroup is of order p^{km-2} and type $m, m, \cdots, m, m-1, m-1$. H_2 is transformed by U according to an operator of order p which transforms one of its operators into a set of generators. Moreover, H_2 contains a characteristic subgroup of order p^{k-2} and type $1, 1, \cdots$ one of whose subgroups of order p is composed of operators invariant under U, and hence the process may be repeated. We have thus the following theorem:

(5.3) *If H is of order $p^{k_1 m + k_2(m-1)}$ and if a set of independent generators of H consists of k_1 operators of order m and k_2 operators of order $m-1$, where $k_1 + k_2 = p - 1$, then the group of isomorphisms of H contains an operator of order p which transforms an operator of H into a set of generators.*

If H is of order p^n but not of the type considered in (5.3), then there are two possibilities: (1) H may have fewer than $p-1$ independent generators, or (2) the independent generators of H may have orders differing from p^m and p^{m-1} either in the number of different orders or in the differences between the orders if they are of just two orders. In the first case we may apply the method of the preceding paragraph and arrive at a quotient group H_r of order p^k and type $2, 2, 1, 1, \cdots$ where $k < p-1$, unless H is itself of type $2, 1, 1, \cdots$

Let us then suppose that H is of order p^{k+2} and type $2, 2, 1, 1, 1, \cdots$ where $k < p-1$. If I contains an operator U of order p which transforms an operator of H into a set of generators, a set of independent generators of H may be chosen so that U transforms the characteristic subgroup of order p^2 composed of p-th powers in H according to the transformation $\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$, which is (4.1) for $k = 2$. Moreover, the generators of H may be selected so that

$$U^{-1}s_1U = s_2,$$
$$U^{-1}s_2U = s_1^{-1}s_2^{2}s_3$$

where s_3 is an operator of order p not in $\{s_1, s_2\}$. If $k > 3$, then $U^{-1}s_3U$ will not be in $\{s_1, s_2, s_3\}$ and may be taken for s_4. Hence we may write U as follows:

$$(5.4)\qquad U=\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ a_1p & a_2p & a_3 & a_4 & \cdots & a_{k-1} & a_k \end{bmatrix}.$$

The last $k-2$ rows correspond to transforms of the $k-2$ independent generators of order p, which explains the factor p in each of the first two elements of the last row. Now if a set of numbers $a_1, a_2, \cdots, a_k$ exists such that U is of order p, the set of numbers $a_3, a_4, \cdots, a_{k'+2}$, $k'+2=k$, is the same as the set $a_1, a_2, \cdots, a_{k'}$ of (4.1) for the same p and $k=k'$. For, first, the quotient group of H with respect to $\{s_1^p, s_2^p\}$ is of order p^k and type $1, 1, \cdots$, and is transformed by U according to an operator obtained from (5.4) by reducing its elements mod p; and, secondly, this operator transforms $s_3=(0, 0, 1, 0, \cdots, 0)$ in the same way $(1, 0, 0, \cdots, 0)$ is transformed by (4.1) when $k=k'$. The numbers $a_3, a_4, \cdots, a_k$ are then given by (4.8). These numbers may then be used in (5.4) to determine a_1 and a_2. The result is that when $k<p-1$ it is impossible to determine a_1 and a_2 so that U is of order p.

Let $r=p-1-k$. Then making use of (4.8) we see that (5.4) takes the form

$$(5.5)\qquad \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ a_1p & a_2p & -\binom{p-1}{r+2} & -\binom{p-2}{r+2} & \cdots & -\binom{r+4}{r+2} & -\binom{r+3}{r+2} \end{bmatrix}.$$

The first k conjugates of $(1, 0, 0, \cdots, 0)$ under (5.5) are successively

$$(5.6)\qquad \begin{matrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 2 & 1 & 0 & \cdots & 0 & 0 \\ -2 & 3 & 2 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -(k-3) & (k-2) & (k-3) & (k-4) & \cdots & 1 & 0 \\ -(k-2) & (k-1) & (k-2) & (k-3) & \cdots & 2 & 1 \end{matrix}.$$

Let us denote the $(k+m)$-th conjugate by

$$(5.7) \qquad a_{m1}, a_{m2}, \cdots, a_{mk}.$$

Let us consider first those numbers a_{mn} for $n \geq 3$, that is, those which denote powers of independent generators of order p. We propose to show that a_{mn} is divisible by $\binom{r+1}{m}$, so that when $m = r + 2$ we have $a_{mn} \equiv 0$, mod p, for $n = 3, 4, \cdots, k$.

From the definition of a_{mn} it follows that

$$(5.8) \qquad a_{mn} = a_{m-1\, n-1} - \binom{r+3+k-n}{r+2} \cdot a_{m-1\, k}$$

for $m \geq 2$ and $n > 2$. The following identity also holds for $m \geq 2$ and $n > 2$:

$$(5.9) \qquad -(m + k - n) a_{mn} = (r - m + 2) a_{m-1\, n} - (r + k - n + 3) a_{m.n+1}.$$

We prove (5.9) by induction, assuming it to be true when m is replaced by $m - 1$, having shown it to be true for $m = 2$.

When $m = 2$, (5.9) becomes

$$(5.10) \qquad -(k + 2 - n) a_{2n} = r a_{1n} - (r + k - n + 3) a_{2n+1}.$$

Using (5.8) to remove a_{2n} and a_{2n+1} and collecting terms we have

$$(n - k - 2) a_{1n-1} + \binom{2-n}{r+2} a_{1k} = (r + n - 2) a_{1n}.$$

This follows from the facts that $(2 - n)\binom{1-n}{r+2} = \binom{2-n}{r+2} \cdot (-n - r)$, and that $k + r + 2 = p + 1 \equiv 1$, mod p. If now we substitute the values of a_{1n-1}, a_{1n}, and a_{1k}, obtained by using the last row of (5.6) in (5.5) we have

$$\begin{aligned}(n - k - 2)\left[k - n + 3 - \binom{3-n}{r+2}\right] + \binom{2-n}{r+2}\left[2 - \binom{r+3}{r+2}\right] \\ = (r + n - 2)\left[k - n + 2 - \binom{2-n}{r+2}\right]\end{aligned}$$

which is readily shown to be true, thus establishing (5.10).

We now suppose (5.9) to hold for all smaller values of m, and by means of (5.8) replace $a_{m-1\, n}$, a_{mn}, and $a_{m\, n+1}$ by expressions containing $a_{m-2\, i}$'s and $a_{m-1\, j}$'s. This gives a congruence which can be written as the sum of the two congruences

$$(5.11) \qquad (r - m + 3) a_{m-2\, n-1} - (r - m + n + 1) a_{m-1\, n-1} \equiv (3 - n) a_{m-1\, n}$$

and

$$\begin{aligned}- a_{m-2\, n-1} - (r - m + 2)\binom{2-n}{r+2} a_{m-2\, k} + (r - m + n + 1)\binom{2-n}{r+2} a_{m-1\, k} \\ \equiv - a_{m-1\, n} - (2 - n)\binom{1-n}{r+2} a_{m-1\, k}.\end{aligned}$$

Now (5.11) is (5.9) in which m has been replaced by $m-1$, and hence the truth of (5.9) depends on the truth of the second congruence. In this we may replace $a_{m-1\,n}-a_{m-2\,n-1}$ by $-\binom{2-n}{r+2}a_{m-2\,k}$, as a result of (5.8). Making this substitution and collecting terms we have

$$-(r-m+3)\binom{2-n}{r+2}\cdot a_{m-2\,k}$$
$$+\left[(r-m+n+1)\binom{2-n}{r+2}+(2-n)\binom{1-n}{r+2}\right]a_{m-1\,k}\equiv 0.$$

The coefficient of $a_{m-1\,k}$ may be simplified and the factor $\binom{2-n}{r+2}$ removed from both terms to give

$$(r-m+3)a_{m-2\,k}\equiv(1-m)a_{m-1\,k}.$$

This congruence is true provided

(5.12) $$a_{mk}\equiv(-1)^m\binom{r+1}{m},$$

for every m.

Hence (5.9) holds for $m=m_1$ provided both (5.9) and (5.12) hold for $m\leq m_1-1$. We compute a_{mk} on the supposition that (5.9) and (5.12) hold for smaller values of m. From (5.9) we have

(5.13) $$-(m-1+k-k+1)a_{m-1\,k-1}$$
$$\equiv(r-m+3)a_{m-2\,k-1}-(r+k-k+4)a_{m-1\,k}.$$

If we compute a_{mk} using (5.8) and take $a_{m-1\,k-1}$ from (5.13) we get

(5.14) $$-m\cdot a_{mk}\equiv(r-m+3)a_{m-2\,k-1}-(r+4)a_{m-1\,k}+m(r+3)a_{m-1\,k}.$$

From (5.8) we have $a_{m-2\,k-1}\equiv a_{m-1\,k}+(r+3)a_{m-2\,k}$. Using this in (5.14) we get

(5.15) $$-ma_{mk}=(r-m+3)[a_{m-1\,k}+(r+3)a_{m-2\,k}]-(r+4)a_{m-1\,k}$$
$$+m(r+3)a_{m-1\,k}.$$

We are assuming (5.12) to hold for $m-i$, and hence (5.15) becomes

(5.16) $$a_{mk}\equiv-\binom{r+2-m}{m}a_{m-1\,k}\equiv(-1)^m\cdot\binom{r+1}{m}.$$

Hence (5.12) holds for $m=m_1$ provided (5.9) holds for $m\leq m_1$ and (5.12) holds for $m\leq m_1-1$. But the latter condition implies the former and therefore (5.12) holds for every m and (5.9) holds for $m\geq 2$, since the former holds when $m=1$.

By means of (5.9) and (5.12) it is easy to show that a_{mi}, $i=3,4,\cdots,k$, is divisible by $\binom{r+1}{m}$ and hence that $a_{r+2\,i}\equiv 0$, $i=3,4,\cdots,k$.

We are now able to compute the first two numbers in the successive conjugates of $(1, 0, 0, \cdots, 0)$. These numbers are congruent, mod p, to the corresponding numbers in the conjugates of $(1, 0, 0, \cdots, 0)$ under (5.5) when a_1 and a_2 are both zero. Hence we may write the $(k+m)$-th pair as $-(k+m-2)+p\,a'_{m1}$, $(k+m-1)+p\,a'_{m2}$. We wish to show that the $(k+r+2)$-th conjugate is $-p+1, p, 0, 0, \cdots, 0$ by showing that $a'_{r+2\,1} \equiv a'_{r+2\,2} \equiv 0$ independently of the values of a_1 and a_2, thus proving that (5.5) is not of order p.

By direct computation we obtain

$$
(5.17)\qquad
\begin{aligned}
a'_{11} &= a_1, & a'_{12} &= a_2;\\
a'_{21} &= a_{1\,k}a_1 - a_2, & a'_{22} &= a_1 + (2 + a_{1\,k})a_2;\\
a'_{31} &= \binom{r-1}{1}\binom{r+2}{2}\cdot a_1 + \binom{r-1}{1}a_2, & a'_{32} &= -\binom{r-1}{1}a_1 + \binom{r-1}{2}a_2.
\end{aligned}
$$

This third pair suggests the general case which is

$$
(5.18)\qquad
\begin{aligned}
a'_{m1} &= (-1)^{m-1}\left[\binom{r-1}{m-2}\binom{r+m-1}{m-1}a_1 + \binom{r-1}{m-2}a_2\right],\\
a'_{m2} &= (-1)^{m}\binom{r-1}{m-2}a_1 + (-1)^{m+1}\binom{r-1}{m-1}a_2.
\end{aligned}
$$

To prove this we compute $a'_{m+1\,1}$ and $a'_{m+1\,2}$ from (5.18), (5.12), and (5.5).

$$a'_{m+1\,1} = \left\{-(-1)^m\binom{r-1}{m-2} + a_{mk}\right\}a_1 - (-1)^{m+1}\binom{r-1}{m-1}a_2.$$

Substituting the value of a_{mk} from (5.12), we have

$$a''_{m+1\,1} = \left\{(-1)^{m+1}\binom{r-1}{m-2} - (-1)^m\binom{r+1}{m}\right\}a_1 + (-1)^{m+2}\binom{r-1}{m-1}a_2$$

This reduces to

$$a'_{m+1\,1} = (-1)^m\left[\binom{r-1}{m-1}\binom{r+m}{m}a_1 + \binom{r-1}{m-1}a_2\right].$$

which is the value of $a'_{m+1\,1}$ given by (5.18) in which m is replaced by $m+1$. The value of $a'_{m+1\,2}$ may be verified in the same manner. Thus (5.18) is established for all values of m.

The $p+1$-st conjugate of $(1, 0, 0, \cdots, 0)$ is $-(p-1)+pa'_{m1}$, $p+pa'_{m2}, a_{m3}, a_{m4}, \cdots, a_{mk}$ where $m = r+2$. Substituting this value for m in (5.18) we have $a'_{r+2\,1} = a'_{r+2\,2} = 0$. Hence the above conjugate is not s_1 and U cannot be of order p.

It is not necessary to compute $a_{r+2\,j}$, $j = 3, 4, \cdots, k$ since it follows from the construction of (5.5) that they are all zero. In fact the intricate computation involving the a_{mn}'s and resulting in (5.8), $\cdots$, (5.16) was necessary only because it is necessary to have (5.12) in order to obtain (5.18).

There will be further use made of these results in what follows. So far we have proved the following theorem:

(5.19) *The group of isomorphisms of the group H of order p^{k+2} and type $2, 2, 1, 1, \cdots$ contains no operator of order p which transforms one of the operators of H into a set of generators unless $k = p - 1$.*

It will be well to notice how $k = p - 1$ affects the preceding considerations. In that case $r = 0$ and m can be at most 2. In (5.18) m must be at least 3. When $m = 2$, a'_{m1} and a'_{m2} are obtained from (5.17). The second pair of equations in (5.17) may be solved for a_1 and a_2 after a'_{m1} and a'_{m2} have been assigned values so that $-(p-1) + pa'_{2\,1} \equiv 1$, and $p + pa'_{2\,2} \equiv 0$, mod p^2. That there exists such a solution follows from (5.3).

Let us suppose now that H has k independent generators of two different orders, k_1 of order p^{m_1} and k_2 of order p^{m_2}. Let us suppose that $m_1 > m_2$ and that $k_1 \geqq 2$. Suppose also that I contains an operator U of order p which transforms an operator of H into a set of generators. The characteristic subgroup H_1 composed of operators of H which are p^{m_2-1}-th powers will have k independent generators, k_1 of order $p^{m_1-m_2+1}$ and k_2 of order p, and this group will be transformed by U according to an operator of order p which transforms one of its operators into a set of generators. The characteristic subgroup H' of order p^{k_1} and type $1, 1, \cdots$ composed of operators which are p^{m_1-1}-th powers in H is also transformed by U according to an operator of order p which transforms one of its operators into a set of generators, and by (4.6) contains a subgroup of order p invariant under U. Let H'_1 be the quotient group of H_1 with respect to this subgroup of order p. H'_1 will have k independent generators, $k_1 - 1$ of order $p^{m_1-m_2+1}$, one of order $p^{m_1-m_2}$, and k_2 of order p. H'_1 is transformed by U according to an operator of order p which transforms one of its operators into a set of generators. This argument may now be repeated and successive quotient groups obtained. The result of a single application is to reduce the number of independent generators of highest order by one, replacing the one by a generator of order $1/p$-th of its order. Let the process be continued to obtain a quotient group H'_i which has 2 generators of order p^2 and $k - 2$ generators of order p. Then the group H'_i is transformed by U according to an operator of order p which transforms one of its operators into a set of generators, which contradicts (5.19) when $k < p - 1$. Hence the operator U does not exist, when $k < p - 1$.

If a set of independent generators of H contains operators of more than two distinct orders, let the number of order p^{m_i} be k_i, and let $m_i > m_{i+1}$. Let U be an operator of order p and of the type in question. Then $k_1 + k_2 + k_3$

$\leq p-1$. If $k_1 = 2$, then the subgroup composed of operators which are p^{m_3}-th powers in H will be of the type just considered, and from (5.19) it follows that U cannot be of order p. If $k_1 = 1$, then the subgroup of order $p^{m_1-m_2}$ contained in a cyclic group of order p^{m_1} is invariant under U and the corresponding quotient group is of the type considered above. In that case the existence of U implies that $k_1 + k_2 + k_3 = p - 1$.

We may summarize the last results as follows:

(5.20) *If the group of isomorphisms of H contains an operator of order p which transforms one of the operators of H into a set of generators and leaves fixed no operator of order greater than p, then a set of independent generators of H contains operators of at most three distinct orders; if there are generators of three distinct orders there is but one of highest order and there are $p-1$ independent generators; if there are operators of two distinct orders there is but one of highest order or there are $p-1$ independent generators.*

It is now necessary to make another computation. We shall prove the following theorem:

(5.21) *The group of isomorphisms of the group H of order p^4 and type 2, 2 contains no operator of order p which transforms an operator of H into a set of generators.*

Suppose such an operator exists and denote it by U. Then generators of H may be chosen so that U takes the form

$$U = \begin{pmatrix} 0 & 1 \\ -1+ap & 2+bp \end{pmatrix}. \tag{5.22}$$

The argument to establish (5.22) is the same as that to establish (5.5). It can be readily verified that the k-th conjugate of (1, 0) is

$$\begin{gathered} -(k-2) + p\left[-\binom{k-1}{2}\binom{k-6}{3}a - \binom{k-1}{3}b\right], \\ k-1+p\left[\binom{k-1}{3}a + \binom{k}{3}b\right]. \end{gathered} \tag{5.23}$$

Since the terms are to be reduced mod p^2, when $k = p+1$ (5.23) becomes $-(p-1)$, p, provided $p > 3$, and hence (5.21) follows.

We have already shown, (5.20), that if H permits an isomorphism of the type in question, then it has independent generators of at most three orders, and $k_1 = 1$. If k_2 is positive, then $m_2 - m_3 = 1$, for by applying the method used to prove (5.20) we should obtain after i steps a quotient group H_i with a set of independent generators, two of order p^{m_2}, $k_1 + k_2 - 2$ of

order p^{m_2-1}, and k_3 of order p^{m_3}, which admitted an isomorphism of order p. Then the group of p^{m_3}-th powers would admit such an isomorphism which would contradict (5.19) unless $p^{m_2-m_3}=p$.

We may then state the following theorem, which includes the last one:

(5.24) *If H admits an isomorphism of the type we are considering, then H is of one of the following types:*

(1) *H has k independent generators of order p, where $k \leq p-1$;*
(2) *H has k_1 independent generators of order p^{m_1} and k_2 of order p^{m_1-1}, where $k_1+k_2=p-1$;*
(3) *H has one independent generator of order p^{m_1}, k_2 of order p^{m_2}, and k_3 of order p^{m_2-1}, where $1+k_2+k_3=p-1$;*
(4) *H has one independent generator of highest order and k_2 of order p where $1+k_2 \leq p-1$.*

We have shown that such isomorphisms exist for groups of type (1) in § 4, and for groups of type (2) at the beginning of the present section. We consider next groups of type (4).

Let H be of order p^{m+k_2} and type $m, 1, 1, \cdots$, and let $k_2+1=k$. Then the quotient group H_1 of H with respect to the group which is composed of p-th powers is of order p^k and type $1, 1, \cdots$. Hence the generators of H may be so chosen that H_1 is transformed by U according to an operator of the form (4.1). The matrix of this operator is obtained from the matrix of U by reducing the elements of the latter according to the modulus p^{m-1}. Moreover, the group composed of operators of H which are p-th powers is invariant under U and contains no operator of order greater than p which is invariant under U. Let the generator of highest order be s_1. Then $U^{-1}s_1^pU=(s_1^p)^{1+ap}$ and since U is of order p we have $U^{-p}s_1^pU^p=(s_1^p)^{(1+ap)^p}=s_1^p$. Hence we have $(1+ap)^p\equiv 1$, mod p^{m-1}, the order of s_1^p. This requires that $ap^2\equiv 0$, mod p^{m-1}, and hence that $m\leq 3$.

If $m=3$ and $k_2=0$, then H is cyclic of order p^3 and the group of isomorphisms of H contains the required operator of order p, viz: $U^{-1}s_1U=s_1^{1+p}$. But if $k_2=1$, no such operator U exists. For, then we should be able to represent U as

$$U=\begin{pmatrix} 1+ap & 1 \\ bp^2 & 1 \end{pmatrix}. \tag{5.25}$$

The $(p+1)$-st conjugate of $(1, 0)$ becomes

$$(1+ap)^p+bp^2[(1+ap)^{p-2}+2(1+ap)^{p-3}$$
$$+\cdots+(p-2)(1+ap)+(p-1)],\ p\cdot$$

which is $1 + ap^2 + bp^2 [p(p-1)/2]$, 0, or $1 + ap^2$, 0. Hence if H is of type (4) and $k_2 = 1$, then the highest order of an operator in H is p^2.

Now suppose that $k_2 > 1$ and $m = 3$. Then generators of H may be so chosen that

$$(5.26)\quad U = \begin{bmatrix} 1+ap & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ bp^2 & -\binom{p-1}{r+1} & -\binom{p-2}{r+1} & \cdots & -\binom{r+3}{r+1} & -\binom{r+3}{r+1} \end{bmatrix}$$

This is established in exactly the same way as (5.5). The conjugates of $1, 0, 0, \cdots, 0$ may be found in the same manner as (5.6), (5.17), and (5.18). The $(k+m)$-th conjugate has for its first element

$$(5.27)\qquad a_{m1} = (1+ap)^{k+m+1} + bp^2 \left[1 - \binom{r+1}{1} + \binom{r+1}{2} - \cdots + (-1)^{m-1}\binom{r+1}{m-1}\right].$$

When $m = r + 2$, this becomes

$$a_{r+2\,1} = (1+ap)^p + bp^2 [(1-1)^{r+1}] = 1 + ap^2.$$

Hence, the $(p+1)$-st conjugate of s_1 is not s_1 and therefore U is not of order p. From this it follows that if H is of type (4), then the generator of highest order is of order p^2.

Now if $m = 2$, then U may be written in the form (5.26) excepting that bp^2 is replaced by bp. Then (5.27) becomes $a_{r+2\,1} \equiv 1$, so that U is of order p. Hence, there exists an operator U of order p in the group of isomorphisms of H when H is of type (4), $m = 2$, and $1 + k_2 \leq p - 1$, and U transforms an operator of highest order in H into a set of generators.

When H is of type (3), its subgroup of p^{m_3}-th powers is of type $m_1 - m_3$, 1, 1, $\cdots$ and has $1 + k_2$ independent generators, and hence is of type (4). Therefore, $m_1 - m_3 = 2$. So the orders of independent generators of H must be p^{m_1}, p^{m_1-1}, and p^{m_1-2} and they must be $p - 1$ in number, with one of highest order. Let H have one generator of order p^m, k_2 of order p^{m-1}, and k_3 of order p^{m-2}, where $1 + k_2 + k_3 = k = p - 1$. Suppose U is of order p and transforms an operator of H into a set of generators. Then the quotient group H_1 of H with respect to the group of p^{m-3}-th powers is of type 3, 2, 2 $\cdots$ 2, 1, 1, $\cdots$ H_1 is transformed by U according to an operator of order p which transforms one of its operators into a set of generators. We shall take this quotient group to be H. The generators of H may be chosen so that U takes the form

$$
(5.28)\qquad
\begin{bmatrix}
1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\
ap & -\binom{p-1}{r_2} & -\binom{p-2}{r_2} & -\binom{p-3}{r_2} & \cdots & -\binom{r_2+1}{r_2} & 1 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 \\
a_1p^2 & a_2p & a_3p & a_4p & \cdots & a_{k_2+1}p & -\binom{p-1}{r_3} & -\binom{p-2}{r_3} & \cdots & -\binom{r_3+2}{r_3} & -\binom{r_3+1}{r_3}
\end{bmatrix}.
$$

We proceed to establish (5.28) in detail. The p-th power of s_1 is not invariant under U, but the p^2-th power is invariant. Then if $U^{-1}s_1U = s_1s_a$, the order of s_a is p^2, and it is not in $\{s_1\}$. Moreover, no power of s_a can be in $\{s_1\}$, for in that case k_2 would be zero, and H would not be of type (3). Hence, we may take s_a for the generator s_2. Then $U^{-1}s_2U$ is an operator of order p^2 which is not in $\{s_1, s_2\}$ and hence may be taken for the operator s_3 in the set of independent generators. We may continue in this manner $U^{-1}s_iU = s_{i+1}$, $i = 2, 3, \cdots, k_2$, until we obtain s_{k_2+1}, the last generator of order p^2. Then $U^{-1}s_{k_2+1}U = s' \cdot s_a$ where s' is an operator of $\{s_1, s_2, \cdots, s_{k_2+1}\}$ and s_a is an operator of order p not in that group. The operator s' is of order p^2. We may select s_a so that $s' = s_1^{x_1}s_2^{x_2}\cdots s_{k_2+1}^{x_{k_2+1}}$ is such that $x_i < p$, $i = 2, 3, \cdots, k_2+1$ and $x_1 < p^2$, for $s_1^{p^2}$ and s_i^p, $i = 2, 3, \cdots, k_2+1$, are of order p and hence may be taken with s_a to give s_{k_2+2}. Hence the elements of the (k_2+1)-th row of (5.28) are all less than p except the first which is ap where a is less than p. The group of p-th powers of H is transformed by U according to an operator whose matrix is the (k_2+1)-rowed

square matrix in the upper left-hand corner of (5.28) whose elements are reduced mod p, which affects none of those elements unless $m = 3$ and then affects only the element ap. Hence we may find the (k_2+1)-th row of (5.28) by methods used to give the last row of (5.26) and the last row of (5.5). This row is

$$(5.29) \qquad ap, -\binom{p-1}{r_2}, -\binom{p-2}{r_2}, \cdots, -\binom{r_2+1}{r_2}, 1, 0, \cdots, 0,$$

where $r_2 = p - 1 - k_2$.

The (k_2+1)-th row introduces s_{k_2+2}, the first of the independent generators of order p. U must transform s_{k_2+2} into an operator not in $\{s_1, s_2, \cdots, s_{k_2+2}\}$ which we may denote by s_{k_2+3}. In general $U^{-1}s_{k_2+i}U = s_{k_2+i+1}$, $i = 3, 4, \cdots, k_3$. Then we wish to indicate the operator $U^{-1}s_{p-1}U$ in such a way as to determine the last row of (5.28). Since s_{p-1} is of order p the first element is a multiple of p^2, and each of the next k_2 is a multiple of p. Each of the last k_3 elements is less than p, since it is the exponent of an operator of order p.

We may determine these last k_3 elements by a consideration of the quotient group H_1 of H with respect to the subgroup of p-th powers. H_1 is of order p^k, $k = 1 + k_2 + k_3$, and type $1, 1, \cdots$, and is transformed by (5.28) according to an operator obtained from (5.28) by reducing each of its elements of mod p. This affects only the first element in the (k_2+1)-th row and the first k_2+1 elements of the last row, each of which becomes zero. We shall designate this new operator by U_1 and consider the conjugates of s_{k_2+2} under U_1. The first k_2+1 elements in the symbol for s_{k_2+2} are zeros and consequently none of its conjugates under U_1 has any of its first k_2+1 elements different from zero. Therefore the argument preceding (4.9) applies and the last row of U_1 is

$$(5.30) \qquad 0, 0, \cdots, 0, -\binom{p-1}{r_3}, -\binom{p-2}{r_3}, \cdots, -\binom{r_3+1}{r_3},$$

where $r_3 = p - 1 - k_3$. The last row of (5.28) is therefore

$$(5.31) \quad a_1p^2, a_2p, a_3p, \cdots, a_{k_2+1}p, -\binom{p-1}{r_3}, -\binom{p-2}{r_3}, \cdots, -\binom{r_3+1}{r_3}.$$

The conjugates of s_1 under U are determined and conditions on $a_1, a_2, \cdots, a_{k_2+1}$ are obtained by requiring U to be of order p. It will be seen that these conditions are inconsistent from which it is inferred that U is not of order p.

The i-th conjugate of s_1 has its first i elements equal to 1 and the rest are zeros, $i = 1, 2, \cdots, k_2+1$. The (k_2+2)-th conjugate is then obtained by adding the first k_2+1 rows of (5.28) and is

$$(5.32) \quad 1 + ap, 1 - \binom{p-1}{r_2}, 1 - \binom{p-2}{r_2}, \cdots, 1 - \binom{r_2+1}{r_2}, 1, 0, \cdots, 0.$$

It will be observed that there are $p-1-r_2=k_2$ elements of the form $1-\binom{r_2+i}{r_2}$ in (5.32), and also that the first k_2 rows of (5.28) have no effect on the last k_3 elements of (k_2+3)-th conjugate of s_1. Let us designate the (k_2+m)-th conjugate of s_1 by

$$(5.33)\qquad a_{m-1\,1},\, a_{m-1\,2},\cdots,\, a_{m-1\,k}.$$

Then $a_{m\,i+1}=a_{m-1\,i}$, for $k_2+1<k_2+m\leqq k=p-1$, and $i\geqq k_2+1$. From this follows

$$(5.34)\qquad \begin{aligned} a_{r_2\,k} &= a_{0\,k_2+1}=1,\\ a_{r_2\,k-1} &= a_{1\,k_2+1}=1-\binom{r_2+1}{r_2}=-\binom{r_2}{1},\\ a_{r_2\,k-2} &= a_{2\,k_2+1}=a_{1\,k_2}-\binom{r_2+1}{r_2}a_{1\,k_2+1}=\binom{r_2}{2},\\ &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ a_{r_2\,k-i} &= a_{i\,k_2+1}=(-1)^i\binom{r_2}{i},\quad 1\leqq i\leqq r_2-2. \end{aligned}$$

To establish the last statement in (5.34) it is necessary to notice (1) that the elements a_{mn}, $n>k_2+1$, are independent of the numbers $a, a_1, a_2,\cdots, a_{k_2+1}$ of (5.28) because s_i, $i>k_2+1$, is of order p and each a, a_i enters with a factor p; and (2) that the elements $a_{m\,k_2+1}$ are the same as the elements a_{mk} given by (5.12) except for multiples of p. Then (5.34) gives $a_{r_2\,n}$ for $n=k_2+2,\, k_2+3,\cdots, k$.

We now seek the elements a_{mn}, $n<k_2+2$. The purpose of the determination of the numbers $a_{r_2\,n}$, $n=1,2,\cdots,k$, is to use them in the determination of the numbers $a_{r_2+2\,n}$ which give the $(p+1)$-th conjugate of s_1. It will be sufficient to determine k_2 of them, as we shall prove. From these we shall get conditions on $a_2, a_3,\cdots, a_{k_2+1}$ which we shall show are inconsistent. Now since s_i, $i=2,3,\cdots,k_2+1$, is of order p^2, the expression for a_{r_2+2n} will be linear in the a_j's and a. We shall for the moment therefore put off a closer determination of a_{mn} and write the k-th conjugate of s_1.

$$(5.35)\qquad \begin{aligned} &a_{r_2\,1},\, a_{r_2\,2},\cdots,\, a_{r_2\,k_2+1},\, (-1)^{r_2-2}\binom{r_2}{r_2-2},\\ &\qquad\qquad (-1)^{r_2-3}\binom{r_2}{r_2-3},\cdots,-\binom{r_2}{1},\, 1. \end{aligned}$$

From this we may compute the $(k+1)$-th conjugate:

$$(5.36)\qquad \begin{aligned} a_{r_2+1\,2} &= a_{r_2\,1}-\binom{p-1}{r_2}a_{r_2\,k_2+1}+a_2p\\ a_{r_2+1\,3} &= a_{r_2\,2}-\binom{p-2}{r_2}a_{r_2\,k_2+1}+a_3p\\ &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ a_{r_2+1\,k_2+1} &= a_{r_2\,k_2}-\binom{r_2+1}{r_2}a_{r_2\,k_2+1}+a_{k_2+1}p. \end{aligned}$$

These with a_{r_2+1} are all that are necessary to determine the $(k+2)$-th conjugate. This element $a_{r_2+1\,k}$ is given by

$$a_{r_2+1\,k} = -\binom{r_2}{1} - \binom{r_3+1}{1} = -(r_2 + r_3 + 1) = -1, \bmod p.$$

Then the $a_{r_2+2\,n}$'s, $n = 2, 3, \cdots, k_2 + 1$, are

$$
\begin{aligned}
a_{r_2+2\,2} &= a_{r_2+1\,1} - \binom{p-1}{r_2}\left[a_{r_2\,k_2} - \binom{r_2+1}{r_2} a_{r_2\,k_2+1} + a_{k_2+1}p\right] - a_2p,\\
a_{r_2+2\,3} &= a_{r_2\,1} - \binom{p-1}{r_2} a_{r_2\,k_2+1} + a_2p\\
&\qquad - \binom{p-2}{r_2}\left[a_{r_2k_2} - \binom{r_2+1}{r_2} a_{r_2\,k_2+1} + a_{k_2+1}p\right] - a_3p,\\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
a_{r_2+2\,k_2+1} &= a_{r_2+1\,k_2} - \binom{r_2+1}{r_2} a_{r_2+1\,k_2+1} - a_{k_2+1}p\\
&= \left[a_{r_2\,k_2-1} - \binom{r_2+2}{r_2} a_{r_2\,k_2+1} + a_{k_2}p\right]\\
&\qquad - \binom{r_2+1}{r_2}\left[a_{r_2\,k_2} - \binom{r_2+1}{r_2} a_{r_2\,k_2+1} + a_{k_2+1}p\right] - a_{k_2+1}p.
\end{aligned}
\tag{5.37}
$$

These numbers must be congruent to zero, mod p^2, since the $(p+1)$-th conjugate of s_1 is s_1. Setting them equal to zero and collecting terms in the a_i's we have k_2 linear equations in the k_2 numbers $a_2, a_3, \cdots, a_{k_2+1}$. The part that is independent of the a_i's is congruent to zero, mod p, since they give the $(k+2)$-th conjugate of the operator corresponding to s_1 in the quotient group of H with respect to its subgroup of p-th powers. We shall examine them more closely later. Let us now consider the matrix of the coefficients of the a_i's in the system (5.37). This matrix is

$$
\begin{bmatrix}
-1 & 0 & 0 & 0 & \cdots & 0 & -\binom{p-1}{r_2}\\
1 & -1 & 0 & 0 & \cdots & 0 & -\binom{p-2}{r_2}\\
0 & 1 & -1 & 0 & \cdots & 0 & -\binom{p-3}{r_2}\\
0 & 0 & 1 & -1 & \cdots & 0 & -\binom{p-4}{r_2}\\
\cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot\\
0 & 0 & 0 & 0 & \cdots & 1 & -\left[\binom{r_2+1}{r_2}+1\right]
\end{bmatrix}.
\tag{5.38}
$$

If we add the rows of this matrix the first $k_2 - 1$ sums are all zeros and the last is

$$
\begin{aligned}
&-\left[1 + \binom{r_2+1}{r_2} + \binom{r_2+2}{r_2} + \cdots + \binom{r_2+k_2}{r_2}\right]\\
&\qquad = -\left[1 + \binom{r_2+1}{1} + \binom{r_2+2}{2} + \cdots + \binom{r_2+k_2}{k_2}\right] = -\binom{r_2+k_2+1}{k_2+1}.
\end{aligned}
$$

Since $r_2 + k_2 + 1 = p$, this sum is also zero. Hence the equations obtained by setting the right-hand sides in (5.37) equal to zero are not independent, or else they are not consistent.

The term independent of the a_i's in each of the equations (5.37) is of the form $c_{mn} + d_{mn}(ap)$. The group of p-th powers in H is transformed by U according to an operator obtained by taking the (k_2+1)-rowed square matrix in the upper left-hand corner of (5.28) and since that group, of type 2, 1, 1, $\cdots$, admits such a transformation the numbers $c_{mn} + d_{mn}(ap)$ must

be congruent to zero for $m = r_2 + 1$, $n > 1$, being the same as those obtained from (5.27). We shall show that the numbers $d_{r_2+1\,n}$ are congruent to zero, mod p, and that the numbers $c_{r_2+1\,n}$ are not congruent to zero, mod p^2. By the methods used to establish (5.18) and (5.19) we may establish

$$c_{m\,k_2+1} = (-1)^m \binom{r_2}{m}, \qquad (m = 1, 2, \cdots, r_2). \tag{5.39}$$

and

$$(m + k_2 + 1 - n) c_{mn} = (1 - n) c_{m\,n+1} - (r_2 - m + 1) c_{m-1\,n}. \tag{5.40}$$

Also we have

$$c_{mn} = c_{m-1\,n-1} - \binom{1-n}{r_2} \cdot c_{m-1\,k_2+1} \tag{5.41}$$

which is analagous to (5.8). Replacing n in (5.41) by $n + 1$, so that the first term on the right-hand side is $c_{m-1\,n}$ and substituting its value in (5.40) we have

$$(m + k_2 + 1 - n) c_{mn} = (m - n - r_2) c_{m\,n+1} - (r_2 - m + 1) \binom{-n}{r_2} c_{m-1\,k_2+1}.$$

This formula with (5.39) allows us to determine the numbers c_{mn}, for a given m, successively beginning at the right with $n = k_2$. We are interested only in the numbers $c_{r_2 n}$, in which case the above formula becomes

$$(-n) c_{r_2\,n} = (-n) c_{r_2\,n+1} - \binom{-n}{r_2} c_{r_2-1\,k_2+1} \tag{5.42}$$

Using (5.39) and (5.42) we obtain

$$\begin{aligned} c_{r_2\,k_2+1} &= (-1)^{r_2} \binom{r_2}{r_2} = (-1)^{r_2} \\ c_{r_2\,k_2} &= (-1)^{r_2} \binom{r_2+1}{r_2} \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ c_{r_2\,n} &= (-1)^{r_2} \binom{-n}{r_2}, \qquad (n = 2, 3, \cdots, k_2 + 1). \end{aligned} \tag{5.43}$$

From these we obtain

$$\begin{aligned} c_{r_2+1\,n} &= c_{r_2\,n-1} - \binom{1-n}{r_2} c_{r_2\,k_2+1} \\ &= (-1)^{r_2} \left[\binom{1-n}{r_2} - \binom{1-n}{r_2} \right] \equiv 0, \text{ mod } p. \end{aligned} \tag{5.44}$$

The formulas (5.39) to (5.44) were obtained for residues, mod p, of the actual coefficients involved in the conjugates of $(1, 0, 0, \cdots, 0)$. The residues are the actual coefficients for the case when H is of order p^{2+k_2} and type $2, 1, 1, \cdots$.

We recall that the numbers c_{mn} are determined by (5.41) starting with the numbers $c_{0n} = 1$, $n = 1, 2, \cdots, k_2 + 1$, and $c_{m1} = 1$, $m = 1, 2, \cdots, r_2$. This last condition is not necessary provided n is allowed to take on negative values in the one before it and in (5.41). The numbers c'_{m1}, thereby obtained,

which take the place of c_{m1}, will all be congruent to 1, mod p; it is understood that in (5.41) $\binom{1-n}{r_2}$ is replaced by $\binom{p+1-n}{r_2}$. If we form a set of numbers c'_{mn}, starting with $c'_{on}=1$ where n may be any positive or negative integer or zero, and using (5.41) modified as described above, the numbers c'_{mn}, $0 \leq m$, $n=1, 2, \cdots, k_2+1$, will be congruent to the numbers c_{mn}. Moreover, the proof that (5.39) and (5.40) hold for the residues of the c's, mod p, shows that (5.39), (5.40), and consequently (5.44), hold for the c''s without reduction.

We shall denote the difference between c_{mn} and c'_{mn} by e_{mn}. We then have

$$
\begin{aligned}
& e_{11}=\binom{p}{r_2} \\
& e_{21}=\binom{p+1}{r_2}-\binom{p}{r_2}\binom{r_2}{1} \\
(5.45)\quad & e_{31}=\binom{p+2}{r_2}-\binom{p+1}{r_2}\cdot\binom{r_2}{1}+\binom{p}{r_2}\binom{r_2}{2} \\
& \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
& e_{r_2 1}=\binom{p+r_2-1}{r_2}-\binom{p+r_2-2}{r_2}\binom{r_2}{1}+\cdots+(-1)^{r_2-1}\binom{p}{r_2}\cdot\binom{r_2-1}{r_2}.
\end{aligned}
$$

Since the e's are all divisible by p and since we are interested only in the residues, mod p^2, of the c's, in any combination of the e's we may reduce the coefficients, mod p. The numbers e_{mn} obviously satisfy the relation

$$(5.46) \qquad e_{mn}=e_{m-1\,n-1}-\binom{1-n}{r_2}e_{m-1\,k_2+1}, \quad 2 \leq n \leq k_2+1.$$

Now the sum of the c's obtained by adding the equations in (5.37) is

$$(5.47) \qquad \sum_{n=2}^{k_2+1} c_{r_2+1\,n}=\sum c'_{r_2+1\,n}+\sum e_{r_2+1\,n}.$$

The first term on the right-hand side is zero, since each of its components satisfies (5.44) without reduction, mod p. This sum is therefore S_{r_2+1}, where

$$(5.48) \qquad S_m=\sum_{n=2}^{k_2+1} e_{mn}.$$

From (5.46) we have

$$
\begin{aligned}
(5.49) \qquad S_m = & e_{m-1\,1}+e_{m-1\,2}+\cdots e_{m-1\,k_2} \\
& -\left[\binom{p-1}{r_2}+\binom{p-2}{r_2}+\cdots+\binom{r_2+1}{r_2}\right]e_{m-1\,k_2+1}.
\end{aligned}
$$

Since we may reduce the coefficient of $e_{m-1\,k_2+1}$, mod p, we have

$$(5.50) \qquad S_m=e_{m-1\,1}+S_{m-1}.$$

By repeated application of (5.50) we obtain

$$(5.51) \qquad S_{r_2+1}=\sum_{m=1}^{r_2} e_{m1}$$

for which we may obtain the values of e_{m1} from (5.45).

If we sum the values of the e_{m1}'s in such a way as to collect terms in $\binom{p+j}{r_2}$, $j = 0, 1, 2, \cdots, r_2$, we have

$$(5.52) \quad S_{r_2+1} = \binom{p+r_2-1}{r_2} + \binom{p+r_2-2}{r_2}\left[1 - \binom{r_2}{1}\right] + \binom{p+r_2-3}{r_2} \cdot \left[1 - \binom{r_2}{1} + \binom{r_2}{2}\right] + \cdots + \binom{p}{r_2} \cdot \left[1 - \binom{r_2}{1} + \binom{r_2}{2} - \cdots + (-1)^{r_2-1}\binom{r_2}{r_2-1}\right],$$

which is

$$(5.53) \quad S_{r_2+1} = \binom{p+r_2-1}{r_2} - \binom{p+r_2-2}{r_2}\binom{r_2-1}{1} + \cdots + (-1)^{r_2-1}\binom{p}{r_2}\binom{r_2-1}{r_2-1}.$$

Each of the numbers $\binom{p+j}{r_2}$ above is divisible by p. We have

$$(5.54) \quad \binom{p+r_2-i}{r_2} = \frac{(p+r_2-i)(p+r_2-i-1)\cdots(p+1)p}{i \quad (i+1) \quad \cdots (r_2-1)r_2}\binom{p-1}{i}$$

which may be written, on isolating the factor p/r_2 and reducing the remaining factors, mod p,

$$(5.55) \quad \binom{p+r_2-i}{r_2} = (-1)^i \, p/r_2 \cdot \binom{r_2-1}{r_2-i}^{-1}.$$

Substituting values from (5.55) in (5.53) we have

$$(5.56) \quad S_{r_2+1} = -p, \bmod p^2.$$

We now consider the numbers d_{mn}. For $n = 1$ the number d_{m1} is $\sum_{i=1}^{m-1} c_{i\,k_2+1}$. Hence,

$$(5.57) \quad d_{r_2\,1} = \left[1 - \binom{r_2}{1} + \binom{r_2}{2} - \cdots + (-1)^{r_2-1}\binom{r_2}{r_2-1}\right] = -(-1)^{r_2}.$$

Since d_{mn} is always multiplied by ap and we are considering powers of operators of order p^2, the residues, mod p, are sufficient. For the remaining d_{mn}'s we have the relation

$$(5.58) \quad d_{mn} = d_{m-1\,n-1} - \binom{1-n}{r_2} d_{m-1\,k_2+1}.$$

The first non-zero $d_{m\,k_2+1}$ appears for $m = k_2 + 1$, and is equal to 1. Thus we have

$$(5.59) \quad d_{mn} = d_{m-1\,n-1} = d_{m-n+1\,1} = \sum_{i=1}^{m-n} c_{i\,k_2+1}, \qquad (m = 2, 3, \cdots, k_2 + 1).$$

The d's of the $(2k_2 + 2)$-th conjugate of $(1, 0, \cdots, 0)$ are

$$(5.60) \quad (-1)^{k_2}\binom{r_2-1}{k_2}, \cdots, -\binom{r_2-1}{3}, \binom{r_2-1}{2}, -\binom{r_2-1}{1}, 1, 0, \cdots, 0.$$

By the method used to derive (5.9) the following relation may be established

$$(5.61) \quad (m + k_2 + 1 - n)\, d_{k_2+1+m\,n} = -(2r_2 - m + 1)\, d_{k_2+1+m-1\,.\,n} + (r_2 + k_2 + 1 - n + 1)\, d_{k_2+1+m\,n+1}.$$

From (5.61), (5.60), and (5.57) we may obtain

$$(5.62) \qquad d_{k_2+1+m\ n} = (-1)^m \binom{2r_2}{m}.$$

From this value of $d_{k_2+1+m\ n}$ and the above relations it can readily be shown that $d_{r_2+1\ n}$ is divisible by p, for $n = 2, 3, \cdots, k_2 + 1$. Each $d_{r_2+1\ n}$ is divisible by $d_{r_2+1\ k_2+1}$ which is $(-1)^{r_2-k_2}\binom{2r_2}{r_2+k_2}$. Since $r_2 > 0$ and $r_2 + k_2 = p - 1$, this last quantity is zero, mod p, unless $r_2 = k_2$, in which case m could not be so great as $k_2 + 1 = r_2 + 1$, and the d's are given by (5.59).

It follows from the above considerations that there are two cases to be considered: (a) $k_2 < (p-1)/2$, in which case the d's are all multiples of p and may be disregarded; and (b), $k_2 \geqq (p-1)/2$, in which case the d's are not all divisible by p and must be considered in examining the augmented matrix obtained from (5.38) by the addition of a column corresponding to the parts of $a_{r_2+1\ n}$ independent of the numbers a_i. If the rank of the augmented matrix is not k_2 then there is a relation connecting the rows; if such a relation exists we have seen that it must state that the sum of the rows is zero. We have just seen that this is true so far as the d's are concerned in case (a); it is likewise true with respect to the d's in case (b). For the sum of the d's given by (5.59) is

$$(5.63) \qquad \sum_{n=0}^{r_2-1} \sum_{i=0}^{n} c_{i\ k_2+1} = \sum \sum (-1)^i \binom{r_2}{i}$$
$$= \sum_{n=0}^{r_2-1} (-1)^n \binom{r_2-1}{n} = (1-1)^{r_2-1} = 0.$$

Therefore the equations (5.37) are inconsistent and the operator U of (5.28) is not of order p. From this it follows that there is no group of type (3) of (5.24) that is not of type (2). We may state the following theorem which supercedes (5.24):

(5.64) *If an abelian group H of order p^n admits an isomorphism U of order p which transforms an operator of H into a set of generators, and if H contains no operator of order p^2 invariant under U, then* (1) *H is of type $1, 1, \cdots$ and order p^k, $k \leqq p-1$*; (2) *H is of type $2, 1, 1, \cdots$ and order p^{k+1}, $k \leqq p-1$; or* (3) *H is of type $m, m, \cdots, m, m-1, m-1, \cdots, m-1$, and has $p-1$ independent generators. Conversely, all groups of types* (1), (2), *and* (3) *admit such isomorphisms.*

In order to complete § 1 it is necessary to find the operators of H left invariant by U when H belongs to the second set of groups of the last theorem or when H belongs to the third set and is of type $2, \cdots, 2, 1, \cdots, 1$. In the latter case U takes the form (5.28) except that the first row and

first column of (5.28) are removed and $k_1 + k_2 = p - 1$. The argument necessary to establish this is exactly like that following (5.28) to determine its last $p - 2$ rows; it depends only on the facts that H and U satisfy (1) and (2) of the introduction. If, as in § 4, we let $s = s_1^{x_1} s_2^{x_2} \cdots s_{p-1}^{x_{p-1}}$ and require that s be transformed into itself by U we get a set of homogeneous congruences analagous to (4.3); the first k_1 are to be taken mod p^2 and the rest mod p. The matrix of coefficients is

$$
(5.65)\qquad \begin{pmatrix}
-1 & 0 & 0\cdots & 0 & & -\binom{p-1}{r_1} & 0\cdots & a_1 p \\
1 & -1 & 0\cdots & 0 & & -\binom{p-2}{r_1} & 0\cdots & a_2 p \\
0 & 1 & -1\cdots & 0 & & -\binom{p-3}{r_1} & 0\cdots & a_3 p \\
0 & 0 & 1\cdots & 0 & & -\binom{p-4}{r_1} & 0\cdots & a_4 p \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0\cdots & -1 & & -\binom{r_1+2}{r_1} & 0\cdots & a_{k_1-1} p \\
0 & 0 & 0\cdots & 1 & -1 & -\binom{r_1+1}{r_1} & 0\cdots & a_{k_1} p \\
0 & 0 & 0\cdots & 0 & 1 & & -1\cdots & -\binom{p-1}{r_2} \\
0 & 0 & 0\cdots & 0 & 0 & & 1\cdots & -\binom{p-2}{r_2} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0\cdots & 0 & 0 & & 0\cdots -1 & -\binom{r_2+1}{r_1}
\end{pmatrix}
$$

It is obvious that if there exists any linear combination of the rows with coefficients $y_1, y_2, \cdots, y_{p-1}$ which gives 0 for each column, then $y_1 = y_2 = \cdots = y_{k_1}$, and the remaining y's are zeros. The sum of the first k_1 elements in the k_1-th column is given in the equation which follows (5.38); it is a multiple of p but not of p^2. The rank of (5.65) is $p - 2$, and hence there is but one cyclic subgroup of H composed of operators invariant under U. The x's constitute a set of coefficients by means of which the columns may be combined to give zero in each row. The value of x_{p-1} may be taken to be 1, provided it is not zero, and the congruences may be solved successively beginning with the last for $x_{p-2}, x_{p-3}, \cdots$. Again as a result of the equation which follows (5.38), the $k_1 + 1$-th equation will give for a value of x_{k_1} a number which is a multiple of p. Then the rest of the x's would be multiples of p and s would be of order p. If x_{p-1} is 0, then the $k_1 + 1$-th equation gives $x_{k_1} \equiv 0$, mod p, and hence the order of s is a multiple of p. Since but one of the above possibilities can be realized, it follows from (4.6) that the second is the one.

We have proved the following theorem which is necessary in the proof of (1.2):

(5.66) *If H has $p - 1$ independent generators of orders p and p^2 every*

operator of order p which transforms an operator of H into a set of generators of H leaves invariant the operators of one and only one subgroup, which is of order p.

When H has $k+1$ generators and is of type $2, 1, 1, \cdots$ U is given by (5. 26) in which a is set equal to zero and bp^2 is replaced by bp. This result follows from the fact that H and U satisfy (1) and (2). From a set of congruences analogous to (4. 3) we obtain immediately the result that x_{k+1} is a multiple of p. Then by solving the congruences successively beginning with the last we find that x_i, $i = 2, 3, \cdots, k+1$ is a multiple of p. Hence the operators in H invariant under U are all in the subgroup generated by s_1^p. Therefore

(5. 67) *If H is of type $2, 1, 1, \cdots$ and is generated by a set of conjugates under U of order p, then the operators of H invariant under U constitute a group of order p.*

There is due a word of explanation of the use of (1. 2) throughout § 5 without the proof having been completed. The proof was complete in § 1 except for the last two theorems above. The proofs of those theorems in no way depends on (1. 2); they involve the consideration of two definite groups. The present arrangement seems more convenient.

URBANA, ILLINOIS.

THE JACOBIAN ALGORITHM FOR PERIODIC CONTINUED FRACTIONS AS DEFINING A CUBIC IRRATIONALITY.

By J. B. Coleman.

In a previous article* we found the conditions under which the characteristic equation for a periodic ternary continued fraction is reducible, so that the number defined is a quadratic irrationality. In that case no restrictions were placed upon the relative magnitudes of the partial quotients, p_i and q_i. Negative as well as positive rational integers were allowed, and the case was discussed in which they were any real numbers. In this paper p_i and q_i are considered as positive rational integers, including 0, with the restrictions that $p_i \gtreqless q_i$, and no q is 0. As customary, and without loss of generality, we take $p_1 \neq 0$. In addition to the above restrictions, the Jacobian algorithm† imposes others of which no account is taken in the following discussion.

It is not difficult to show, by direct expansion, that the characteristic equation is irreducible for $k = 1, 2, 3, 4$,‡ where k represents the number of pairs of partial quotients in a period.

The characteristic equation is of the form, $\rho^3 - M\rho^2 + N\rho - 1 = 0$, in which M is a positive rational integer and N is a rational integer, for the Jacobian algorithm. The necessary and sufficient conditions for reducibility are that $\rho = \pm 1$, i. e., that $M = N$ or $M = -N - 2$. If N be positive, $M = N$ is the condition for reducibility, and if N be negative, $M = |N| - 2$ is the condition. In any case, to prove irreducibility, it will be sufficient to show that, in general, $M > |N|$.

In our article cited above, M and N were found, each as the sum of three continuants. These six continuants may be easily expressed as of two types. By $C_k{}^i$ is indicated the type;

$$(1)\qquad \begin{vmatrix} q_i & 1 & & & & & & \\ -p_{i+1} & q_{i+1} & 1 & & & & & \\ 1 & -p_{i+2} & q_{i+2} & 1 & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & 1 & -p_{k-2} & q_{k-2} & 1 & & \\ & & & 1 & -p_{k-1} & q_{k-1} & 1 \\ & & & & 1 & -p_k & q_k \end{vmatrix}$$

* J. B. Coleman, *American Journal of Mathematics*, Vol. 52 (1930), p. 835.

† C. G. J. Jacobi, *Werke*, Vol. 6, p. 385.

‡ O. Perron, *Mathematische Annalen*, Vol. 64, p. 1.

By $F_k{}^i$ is indicated the type;

$$
(2) \qquad \begin{vmatrix}
-p_i & q_i & 1 & & & & & \\
1 & -p_{i+1} & q_{i+1} & 1 & & & & \\
& 1 & -p_{i+2} & q_{i+2} & 1 & & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
& & & 1 & -p_{k-2} & q_{k-2} & 1 \\
& & & & 1 & -p_{k-1} & q_{k-1} \\
& & & & & 1 & -p_k
\end{vmatrix}
$$

In both (1) and (2), $k \geqq i$, and all elements not otherwise designated are 0.

Making use of the above notation, the results previously obtained give

$$
(3) \qquad M = C_k{}^1 + p_1 C^2{}_{k-1} + C^2{}_{k-2} + C^3{}_{k-1}
$$
$$
(C_k{}^i = 1 \text{ if } i = k+1,\ C_k{}^i = 0 \text{ if } i \geqq k+2),
$$

$$
(4) \qquad N = F_k{}^1 - q_k F^2{}_{k-1} + F^2{}_{k-2} + F^3{}_{k-1}
$$
$$
(F_k{}^i = 1 \text{ if } i = k+1,\ F_k{}^i = 0 \text{ if } i \geqq k+2).
$$

To show that $M > |N|$ in general when $k \geqq 5$, we make use of four recursion formulae and five theorems connected with them. Expanding (1) on the last row gives

$$
(5) \qquad C_k{}^i = q_k C^i{}_{k-1} + p_k C^i{}_{k-2} + C^i{}_{k-3}.
$$

Expanding (1) on the first column gives

$$
(6) \qquad C_k{}^i = q_i C_k{}^{i+1} + p_{i+1} C_k{}^{i+2} + C_k{}^{i+3}.
$$

Expanding (2) on the last column gives

$$
(7) \qquad F_k{}^i = -p_k F^i{}_{k-1} - q_{k-1} F^i{}_{k-2} + F^i{}_{k-3}.
$$

Expanding (2) on the first row gives

$$
(8) \qquad F_k{}^i = -p_i F_k{}^{i+1} - q_i F_k{}^{i+2} + F_k{}^{i+3}.
$$

In these recursion formulae the same conditions apply as in (3) and (4), when $i > k$.

THEOREM I. *In the expansion* (7), $F_j{}^i$ *cannot have the same sign for more than two successive values of* j.

If $F_j{}^i$, $F^i{}_{j-i}$, $F^i{}_{j-2}$ all have the same sign, then, since by (7), $-p_j F^i{}_{j-i} - q_{j-i} F^i{}_{j-2}$ must be opposite in sign to $F_j{}^i$, it follows that

$$
| F^i{}_{j-3} | > | -p_j F^i{}_{j-1} - q_{j-1} F^i{}_{j-2} |
$$

and the sign of $F^i{}_{j-3}$ must be the same as that of $F_j{}^i$. This shows that if

F_j^i have the same sign for three successive values of j, it will have the same sign also for the next preceding value of j. However for the three smallest possible values of j, $i-1$, i, and $i+1$, it is impossible for F_j^i to have three like signs. Hence, by induction, it will never be possible for three successive terms to have the same signs.

THEOREM II. *In the expansion* (8), *it is impossible for F_k^i to have the same sign for more than two successive values of i.*

The proof is similar to that of Theorem I.

Note. In the course of the proofs of the following theorems many special cases would arise for small values of k, particularly when $k=1, 2, 3$. For the sake of economy we do not always specify these, because they do not affect the generality of the results.

THEOREM III. $C_k^i > |F_k^i|$ *when* $k > i+1$. ($C_k^i \geqq |F_k^i|$ *when* $k = i, i+1$).

The proof is by induction. Assume

(9a) $$C_j^i > |F_j^i| \quad \text{when} \quad j = k-1, k-2, \quad \text{and}$$

(9b) $$C_j^i \geqq |F_j^i| \quad \text{when} \quad j = k-3, k-4.$$

By (5)

(5a) $$C_j^i = q_j C^i_{j-1} + p_j C^i_{j-2} + C^i_{j-3}.$$

By (7)

(7a) $$F_j^i = -p_j F^i_{j-1} - q_{j-1} F^i_{j-2} + F^i_{j-3}.$$

If F^i_{j-1} and F^i_{j-2} have opposite signs, either

(10a) $$|F_j^i| \gtreqless |-q_{j-1}F^i_{j-2} + F^i_{j-3}| \quad \text{and}$$

(10b) $$|F_j^i| \gtreqless |-p_j F^i_{j-1}|, \quad \text{or}$$

(11a) $$|F_j^i| \gtreqless |-p_j F^i_{j-1} + F^i_{j-3}| \quad \text{and}$$

(11b) $$|F_j^i| \gtreqless |-q_{j-1} F^i_{j-2}|.$$

In (10a) using the fact from (5) that $C^i_{j-1} > q_{j-1} C^i_{j-2}$, from (9a) and (5) it follows that $C_j^i > |F_j^i|$. In (11a), since $p_j \gtreqless q_j$ by the algorithm, from (5), (9a) and (9b) it follows that $C_j^i > |F_j^i|$. The proof for (10b) is included in (11a), and that for (11b) is included in (10a).

If F^i_{j-1} and F^i_{j-2} have the same sign, by Theorem I, both F_j^i and F^i_{j-3}

must be of the opposite sign. Expanding $F^i{}_{j-1}$ by (5) and substituting in (7a)

$$(12)\qquad F_j{}^i = p_jp_{j-1}F^i{}_{j-2} + p_jq_{j-2}F^i{}_{j-3} - p_jF^i{}_{j-4} - q_{j-1}F^i{}_{j-2} + F^i{}_{j-3}.$$

Expanding $C^i{}_{j-1}$ by (5) and substituting in (5a) we have

$$(13)\qquad C_j{}^i = q_jq_{j-1}C^i{}_{j-2} + q_jp_{j-1}C^i{}_{j-3} + q_jC^i{}_{j-4} + p_jC^i{}_{j-2} + C^i{}_{j-3}.$$

Since $F^i{}_{j-2}$ and $F_j{}^i$ have opposite signs, from (12) may be written

$$(14)\qquad | F_j{}^i | \lesseqgtr | p_jq_{j-2}F^i{}_{j-3} - p_jF^i{}_{j-4} - q_{j-1}F^i{}_{j-2} + F^i{}_{j-3} |.$$

We now compare the terms of (13) and (14), using (9a), (9b) and the algorithm with respect to the values of p_j and q_j.

$$\begin{aligned} p_jC^i{}_{j-2} &> p_jq_{j-2}C^i{}_{j-3} \quad \text{from (5)} \\ &> | p_jq_{j-2}F^i{}_{j-3} |, \\ q_jC^i{}_{j-4} &\geqq | p_jF^i{}_{j-4} |, \\ q_jq_{j-1}C^i{}_{j-2} &\geqq | q_{j-1}F^i{}_{j-2} | \quad \text{and} \\ C^i{}_{j-3} &\geqq | F^i{}_{j-3} |. \end{aligned}$$

Hence $C_j{}^i > | F_j{}^i |$.

By direct expansion it is found that $C_j{}^i \geqq | F_j{}^i |$ when $j = i, i+1$ and $C_j{}^i > | F_j{}^i |$ when $j = i+2, i+3$, hence from the above argument $C_j{}^i > | F_j{}^i |$ for $j = i+4$. By induction it will be true for all values of $j > i+1$.

THEOREM IV. *If $F_k{}^i$ and $F^i{}_{k-1}$ have the same sign then* $| F_k{}^i | < C^i{}_{k-1}$.

By Theorem I the sign of $F^i{}_{k-2}$ is opposite to that of $F_k{}^i$. If now the sign of $F^i{}_{k-3}$ is also opposite to that of $F_k{}^i$, then

$$| F_k{}^i | \leqq | -q_{k-1}F^i{}_{k-2} | < C^i{}_{k-1}, \quad \text{from III and (5).}$$

If now F_{k-3} has the same sign as $F_k{}^i$ and $p_{k-1} \neq 0$, then from (7)

$$(15)\qquad | F_k{}^i | \leqq | -q_{k-1}F^i{}_{k-2} + F^i{}_{k-3} |.$$

Comparing this with the expansion for $C^i{}_{k-1}$ by (5) we have by the use of III the result desired.

Next let $F^i{}_{k-3}$ have the same sign as $F_k{}^i$ and $p_{k-1} = 0$. By virtue of these conditions it is evident that from the expansion of $F^i{}_{k-1}$ by (7),

$$| q_{k-2}F^i{}_{k-3} | < | F^i{}_{k-4} | \quad \text{or} \quad | F^i{}_{k-3} | < | F^i{}_{k-4} |.$$

From (15)

$$| F_k{}^i | < | -q_{k-1}F^i{}_{k-2} | + | F^i{}_{k-4} |.$$

In this case, by (5) we have $C^i{}_{k-1} = q_{k-1}C^i{}_{k-2} + C^i{}_{k-4}$, so that by III, $C^i{}_{k-1} > | F_k{}^i |$.

THEOREM V. *$C_j^i > | F_k^h |$, when $h \geqq i, k \gtreqqless j, i+1 < j$. (There may be equality when $i+1, i=j$, the other conditions being the same).*

Since all the terms in the expansion of C_j^i by (5) and (6) are positive, it follows that $C_j^i > C_k^h$ if either $h > i$ or $k < j$. This in connection with III proves the theorem.

To show that $M > | N |$ in general when $k \geqq 5$.

By III the sum of the last two terms in (3) is equal to or greater than the sum of the last two terms in (4). The general result then depends upon proving that

(16) $$C_k^1 + p_1 C^2_{k-1} > | F_k^1 - q_k F^2_{k-1} |.$$

(a) Proof of (16) when $p_1 \geqq q_k$.

Regardless of the signs of F_k^1 and F^2_{k-1}, it is evident from III that (16) is true if $p_1 \geqq q_k$. This leaves unproved the case where $q_k > p_1$, and since by the algorithm $p_1 \geqq 1$, in this case q_k must be greater than 1.

(b) Proof of (16) when the signs of F_k^1 and of F^2_{k-1}, positive or negative, are the same.

From (16) either

(17) $$| F_k^1 - q_k F^2_{k-1} | \gtreqqless | F_k^1 |, \text{ or}$$

(18) $$\gtreqqless | q_k F^2_{k-1} |.$$

In the case of (17) it is evident from III that (16) is true.

In the case of (18), since by (5) $C_k^1 > q_k C^1_{k-1}$, it is evident that (16) is true.

This leaves unproved the case where F^2_{k-1} is opposite in sign to F_k^1.*

(c) Proof of (16) when F^1_{k-1} is s.

By IV, $| F_k^1 | < C^1_{k-1}$. From this and III

(19) $$| F_k^1 - q_k F^2_{k-1} | < C^1_{k-1} + q_k C^2_{k-1}.$$

By (5) $C_k^1 > q_k C^1_{k-1}$, and by the algorithm $p_1 \geqq 1$, hence

(20) $$C_k^1 + p_1 C^2_{k-1} > q_k C^1_{k-1} + C^2_{k-1}.$$

Subtracting the right side of (19) from the right side of (20) gives $(q_k - 1)(C^1_{k-1} - C^2_{k-2})$. This expression is equal to or greater than 0, since $C^1_{k-1} > C^2_{k-2}$ and $q_k \geqq 1$. This proves (16) when F^1_{k-1} has the sign s.

(d) Proof of (16) when $q_1 \gtreqqless p_2$.

* In this discussion the relative signs of the terms are considered so frequently that we indicate the sign of a term by s, and the opposite sign by o. Throughout the discussion F_k^1 is given the sign s.

Expand F_k^1 by (8) and F_k^2 by (7), then substitute from the second expansion in the first. This gives for the right side of (16)

$$(21)\qquad F_k^1 - q_kF^2_{k-1} \\ = p_1p_kF^2_{k-1} + p_1q_{k-1}F^2_{k-2} - p_1F^2_{k-3} - q_1F_k^3 + F_k^4 - q_kF^2_{k-1}.$$

In the same way expand C_k^1 by (6) and C_k^2 by (7), then substitute from the second expansion in the first. Also expand C^2_{k-1} by (5). This gives for the left side of (16)

$$(22)\qquad q_1q_kC^2_{k-1} + q_1p_kC^2_{k-2} + q_1C^2_{k-3} \\ + p_2C_k^3 + C_k^4 + p_1(q_{k-1}C^2_{k-2} + p_{k-1}C^2_{k-3} + C^2_{k-4}).$$

The sign of (21) is s and as a result of (b) we take that of F^2_{k-1} as o, so that the first term of the expansion may be neglected in considering maximum absolute value. Comparing the terms of (21) and (22), making use of III and the algorithm, we get the following;

$$|\, p_1q_{k-1}F^2_{k-2} \,| \lesseqgtr p_1q_{k-1}C^2_{k-2}; \quad |\, p_1F^2_{k-3} \,| \lesseqgtr q_1C^2_{k-3}; \quad |\, F_k^4 \,| \lesseqgtr C_k^4$$

and $|\, q_kF^2_{k-1} \,| < q_1q_kC^2_{k-1}$. Hence if $|\, q_1F_k^3 \,| \lesseqgtr p_2C_k^3$ then (16) must be satisfied. By III this is true if $q_1 \lesseqgtr p_2$. The case where $p_2 < q_1$ remains to be proved.

(e) Proof of (16) if F^1_{k-2} is s.

Expand F_k^1 by (7) and F^1_{k-1} by (8), then substitute from the second expansion into the first. This gives for the right side of (16)

$$(23)\qquad F_k^1 - q_kF^2_{k-1} \\ = p_1p_kF^2_{k-1} + q_1p_kF^3_{k-1} - p_kF^4_{k-1} - q_{k-1}F^1_{k-2} + F^1_{k-3} - q_kF^2_{k-1}.$$

From (b) we take F^2_{k-1} as o, so that the first term on the right of (23) may be neglected. By hypothesis F^1_{k-2} is s, so that the term involving it will decrease the absolute value of (23) unless it is 0. Neglecting these two terms and expanding F^2_{k-1} by (8), (23) reduces to

$$(24)\qquad |\, F_k^1 - q_kF^2_{k-1} \,| \\ \lesseqgtr |\, q_1p_kF^3_{k-1} - p_kF^4_{k-1} + F^1_{k-3} + q_k(p_2F^3_{k-1} + q_2F^4_{k-1} - F^5_{k-1}) \,|.$$

Expand C_k^1 by (5) and C^1_{k-1} by (7), then substitute the value from the second expansion into the first. Also expanding C^2_{k-1} by (6) in this expression gives

$$(25)\qquad C_k^1 = q_1q_k(q_2C^3_{k-1} + p_3C^4_{k-1} + C^5_{k-1}) \\ + p_2q_kC^3_{k-1} + q_kC^4_{k-1} + p_kC^1_{k-2} + C^1_{k-3}.$$

Comparing (24) and (25), using III and the algorithm,

$$(26) \quad |F^1_{k-3}| \gtreqless C^1_{k-3}; \quad |p_2q_kF^3_{k-1}| \gtreqless p_2q_kC^3_{k-1}; \quad |q_kF^5_{k-1}| \gtreqless q_1q_kC^5_{k-1}.$$

We now consider two cases;

(e_1) When F^4_{k-1} is o. (e_2) When F^4_{k-1} is s.

In (e_1) the term $q_kq_2F_{k-1}$ may be neglected in considering the maximum absolute value of (24). For convenience we will also neglect some terms in (25). These together with the relations (26) give from (24) and (25)

$$\begin{aligned} C_k^1 - |F_k^1 - q_kF^2_{k-1}| &\geqq q_1q_2q_kC^3_{k-1} - |q_1p_kF^3_{k-1}| + q_kC^4_{k-1} - |p_kF^4_{k-1}| \\ &\geqq 0, \text{ by the algorithm and III.} \end{aligned}$$

Hence (16) is true in this case.

In (e_2) the term $p_kF^4_{k-1}$ may be neglected in (24). This in connection with (26), (24) and (25) gives

$$\begin{aligned} C_k^1 - |F_k^1 - q_kF^2_{k-1}| &\geqq q_1q_2q_kC^3_{k-1} - |q_1p_kF^3_{k-1}| + q_kC^4_{k-1} - |q_2q_kF^4_{k-1}| \\ &\geqq q_1q_2q_kC^3_{k-1} - q_1q_kC^3_{k-1} + q_kC^4_{k-1} - q_2q_kC^4_{k-1} \\ &\geqq q_1q_k(q_2-1)C^3_{k-1} + q_k(1-q_2)C^4_{k-1} \geqq 0. \end{aligned}$$

The last results follow from the algorithm, III and (5). This proves (16) for this case, leaving the case where F^1_{k-2} is o.

(f) Proof of (16) when F^1_{k-2} is o.

From (c) we need consider only the case when F^1_{k-1} is o, and since F^1_{k-2} is also o, then by IV, $|F^1_{k-1}| < C^1_{k-2}$. Then by (7)

$$|F_k^1 - q_kF^2_{k-1}| < p_kC^1_{k-2} + |-q_{k-1}F^1_{k-2} + F^1_{k-3} - q_kF^2_{k-1}|.$$

Subtracting this from (5), and using III, gives

$$C_k^1 - |F_k^1 - q_kF^2_{k-1}| > q_kC^1_{k-1} - |q_{k-1}F^1_{k-2}| - |q_kF^2_{k-1}|.$$

In this case (16) is true if the above expression is equal to, or greater than 0. That this is so will now be shown by proving that the first member on the right is more than double either of the next two.

After (a) we need consider only the case when $q_k > 1$. In this case, by (5) and III, $q_kC^1_{k-1} > q_kq_{k-1}C^1_{k-2} > 2\,|q_{k-1}F^1_{k-2}|$.

Next we prove that $q_kC^1_{k-1} \geqq 2\,|q_kF^2_{k-1}|$ or $C^1_{k-1} \geqq 2\,|F^2_{k-1}|$. By (6)

$$(27) \qquad C^1_{k-1} = q_1C^2_{k-1} + p_2C^3_{k-1} + C^4_{k-1}.$$

From (27) and III it is evident that $C^1_{k-1} > 2\,|F^2_{k-1}|$, if $q_1 \geqq 2$. It is to be shown that this is true when $q_1 = 1$. In this case $p_1 = 1$ and after (d) we may take $p_2 = 0$. Under these conditions from (8)

$$(28) \qquad F^2{}_{k-1} = -q_2 F^4{}_{k-1} + F^5{}_{k-1}$$

If $F^4{}_{k-1}$ is o, then since $F^2{}_{k-1}$ is o

$$|F^2{}_{k-1}| < |F^5{}_{k-1}| < C^5{}_{k-1}, \quad \text{by III.}$$

From (6) it is readily seen that $C^1{}_{k-1} > 2C^4{}_{k-1}$ and hence $> 2C^5{}_{k-1}$, so that under the conditions stated $C^1{}_{k-1} > 2\,|F^2{}_{k-1}|$.

Lastly let $F^4{}_{k-1}$ be s. Then $F^3{}_{k-1}$ is to be taken as s, by II, since as a consequence of (b) and (c), $F^1{}_{k-1}$ and $F^2{}_{k-1}$ are both to be taken as of sign o. Under the conditions we have from (8) $F^1{}_{k-1} = -F^2{}_{k-1} - F^3{}_{k-1} + F^4{}_{k-1}$. As a result of the signs of these terms

$$(29) \qquad |F^2{}_{k-1}| + |F^4{}_{k-1}| < |F^3{}_{k-1}|.$$

From (29), either $|F^2{}_{k-1}| \gtreqless \frac{1}{2}\,|F^3{}_{k-1}|$ or $|F^4{}_{k-1}| \gtreqless \frac{1}{2}\,|F^3{}_{k-1}|$. In the first of these alternatives, since $|F^3{}_{k-1}| < C^1{}_{k-1}$ by V, $C^1{}_{k-1} > 2\,|F^2{}_{k-1}|$. In the second alternative

$$(30) \qquad |F^4{}_{k-1}| \gtreqless \tfrac{1}{2}\,|F^3{}_{k-1}| \gtreqless \tfrac{1}{2}\,C^3{}_{k-1}, \quad \text{by III.}$$

Since $p_1 = q_1 = 1$, and $p_2 = 0$, from (6)

$$(31) \qquad C^1{}_{k-1} = C^2{}_{k-1} + C^4{}_{k-1} = q_2 C^3{}_{k-1} + p_3 C^4{}_{k-1} + C^5{}_{k-1} + C^4{}_{k-1}.$$

Substituting from (30) in (28) gives

$$(32) \qquad |F^2{}_{k-1}| \gtreqless \tfrac{1}{2}\,q_2 C^3{}_{k-1} + |F^5{}_{k-1}|.$$

Since by V, $C^4{}_{k-1} > |F^5{}_{k-1}|$, it is seen at once from (31) and (32) that $C^1{}_{k-1} > 2\,|F^2{}_{k-1}|$.

This completes the proof of (16) when $F^1{}_{k-2}$ has the sign o, under all conditions for which it was not proved in (a), (b), (c) and (d). In (e), (16) was shown to hold when $F^1{}_{k-2}$ had the sign s, likewise under all conditions not previously proved. This proves in general that the characteristic equation for Jacobian algorithm is irreducible.

The following numerical example will serve to illustrate comparative values of M and N in a typical case where six pairs of partial quotients are taken in the period.

Given $(\overline{1,2;\,4,5;\,2,3;\,3,4;\,2,2;\,5,6};\cdots)$.

By the recursion formulae

$$C_6{}^1 = 4639,\quad C_5{}^2 = 207,\quad C_4{}^2 = 84,\quad C_5{}^3 = 37,$$
$$F_6{}^1 = -72,\quad F_5{}^2 = -24,\quad F_4{}^2 = 4,\quad F_5{}^3 = 3.$$

By (3), $M = 4967$, and by (4), $N = 79$, so that the characteristic equation is

$$\rho^3 - 4967\rho^2 + 79\rho - 1 = 0.$$

MINIMUM DECOMPOSITIONS INTO N-TH POWERS.

By L. E. Dickson.

1. Let $a = 2^n$, $b = 3^n$. Consider all the decompositions $x + ya + zb$ of a given integer i in which x, y, z are integers $\geqq 0$. The case in which $x + y + z$ is the minimum yields the minimum decomposition. We shall find the minimum decompositions of all integers i.

All decompositions of all integers may be exhibited in a highly condensed table whose successive columns involve the successive multiples of b. Down to a certain point every column has a minimum decomposition, while after that point no column has a minimum decomposition. This point of division is a complicated function of n which is by no means monotonic. This function is evaluated for $n \leqq 36$, a limit beyond the needs of applications to Waring's problem. Except for this point, the theory is developed for a general n and is remarkably simple.

2. We employ the following quotients and remainders:

$$b = qa + r \qquad (0 \leqq r < a), \tag{1}$$

$$a = Q(a - r) + R \qquad (0 \leqq R < a - r). \tag{2}$$

Since r is odd, $r > 0$. Since $a - r$ is odd and hence is not a factor of $a = 2^n$, we have $R > 0$.

Theorem 1. *If* $n = 2$, $a = q + r + 1$. *If* $n = 3$, $a = q + r + 2$. *If* $n \geqq 4$, $a > q + r + 3$.

We have $a = q + r + 10$ if $n = 4$. If $n \geqq 5$,

$$(4/3)^n - 9(2/3)^n > 3,$$

since it holds for $n = 5$ and since the first power increases with n and the second power decreases. By (1), $(3/2)^n > q$. Hence

$$\tfrac{1}{3} 2^n > 3 + (3/2)^n > q + 3.$$

This proves the theorem if $\tfrac{2}{3} 2^n \geqq r$. Henceforth let $3r > 2 \cdot 2^n$.

To proceed by induction on n, $n \geqq 4$, assume the theorem with n replaced by $n - 1$. Then

$$3^{n-1} = q_1 2^{n-1} + r_1, \quad 0 < r_1 < 2^{n-1}, \quad 2^{n-1} \geqq q_1 + r_1 + 2, \quad 3r_1 > 2 \cdot 2^{n-1}.$$

Case $q_1 = \text{even} = 2k$. We have $3r_1 = 2^n + \rho$, $\rho > 0$, $6r_1 < 3 \cdot 2^n$, whence $2\rho < 2^n$. Also, $3^n = 3k \cdot 2^n + 3r_1 = (1 + 3k)2^n + \rho$, $0 < \rho < \frac{1}{2} 2^n$. Comparison with (1) gives $r = \rho$, $q = 1 + 3k$. Thus

$$\begin{aligned} 2 \cdot 2^n > 4q_1 + 4r_1 + 8 &= 8k + 4r_1 + 8 \\ &\geqq 1 + 3k + 3r_1 + 8 = 1 + 3k + 2^n + r + 8. \end{aligned}$$

Subtracting 2^n, we get $2^n > q + r + 8$ if $n \geqq 4$.

Case $q_1 = \text{odd} = 2k + 1$. Then

$$3^n = 3(2k + 1)2^{n-1} + 3r_1 = (3k + 1)2^n + 2^{n-1} + 3r_1.$$

Since $3r_1 > 2 \cdot 2^{n-1}$ and $3r_1 < 3 \cdot 2^{n-1}$, we may write $3r_1 = 2^{n-1} + d$, $0 < d < 2^n$. Hence $3^n = (3k + 2)2^n + d$. Comparison with (1) yields $r = d$, $q = 3k + 2$. But

$$2^n > 2(q_1 + r_1 + 2) = 4k + 2 + 2r_1 + 4 > 10/3 + 2k + 2r_1.$$

Multiplication by 3/2 gives

$$a = 2^n > 5 + 3k + 3r_1 - 2^{n-1} = q + d + 3 = q + r + 3.$$

3. The example $n = 7$ will clarify the later theory. Any integer can evidently be expressed as the sum of a number $ma + kb$ and a number chosen from $0, 1, \cdots, 127 = a - 1$. These sums are written in the last column of the following tablette:

$44+(m+68)a+(k-4)b$	$33+(m+51)a+(k-3)b$	$22+(m+34)a+(k-2)b$	$11+(m+17)a+(k-1)b$	$[ma+kb]$
. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
$[(m+69a)+(k-4)b]$	$117+$	$106+$	$95+$	$84+$
. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
$11+$	$[(m+52)a+(k-3)b]$	$117+$	$106+$	$95+$
. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
$22+$	$11+$	$[(m+35)a+(k-2)b]$	$117+$	$106+$
. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
$33+$	$22+$	$11+$	$[(m+18)a+(k-1)b]$	$117+$
. . . . .	. . . . .	. . . . .	. . . . .	. . . . .

The numbers in any line are equal since $b = 17a + 11$. The upper row of dots takes the place of 83 lines obtained by adding $1, 2, \cdots, 83$ in turn to the first line. To the last four rows of dots we add $1, 2, \cdots, 10$. Hence there are $5 + 83 + 4 \times 10 = 128 = a$ rows. The sum of the coefficients ($\geqq 0$) of 1, a, b in a decomposition is called its *weight*. The weight of a number in [ ] will be proved to be less than the weights of the remaining entries of the same line.

If we continue the tablette to the left of the fifth column, we shall prove that the weight of every number in the annexed columns exceeds the weight of the number in [ ] in the same line of the original tablette. Thus any number in [ ] is a minimum decomposition provided the number in the same line and right-hand column is $< 4^n$.

4. The general theory is based on

$$(3)\qquad F(h,i)=(m+hQq+hQ+iq+i-h)a+(k-hQ-i)b.$$

In the table for $n=7$ the quantities in [ ] in the successive columns are $F(3,1)$, $F(2,1)$, $F(1,1)$, $F(0,1)$, $F(0,0)$. The last two columns form the strip 0; while the first three columns form the strips 3, 2, 1, respectively. In the table for any n, the columns containing the $F(h,i)$ with a fixed h and varying i form the strip h, and the $F(h,i)$ having the least i is called the top F of the strip h. Thus for $n=7$, $F(0,0)$ is the top F of the strip 0. We next prove

$$(4)\qquad F(h,i)-ma-kb=\Delta(h,i),\qquad \Delta(h,i)\equiv i(a-r)-hR.$$

Transpose $(-hQ-i)b$ and replace b by its value (1); cancellations yield the product of (2) by h. Next,

$$(5)\qquad F(H,I)-F(H-h,I-i)=F(h,i)-ma-kb=\Delta(h,i).$$

The case $h=0$, $i=j$ of (5) gives

$$(6)\ F(H,I)=j(a-r)+F(H,I-j),\ F(h,j+i)=j(a-r)+F(h,i).$$

THEOREM 2. *Let $F(h,i)$ be the top F of the strip h. Let g be the least integer such that*

$$(7)\qquad (h+1)R-1<g(a-r).$$

Then $F(h+1,g)$ is the top F of the strip $h+1$, while $F(h,Q+g-1)$ is the bottom F of strip h. Also, $\Delta=i(a-r)-hR$ is $\geqq 0$ and $<a-r$. Finally, $g\geqq i$.

The theorem is true if $h=0$. Then F_{00} is the top and F_{0Q} is the bottom F of strip 0 by (2), while F_{11} is the top F of strip 1. Also, $g=1$.

We assume the theorem for a fixed h and prove it true when h is replaced by $h+1$. By (4) and (6_2) with $j=Q+g-1-i$,

$$F(h,Q+g-1)=ma+kb+D,\qquad D=(Q+g-1-i)(a-r)+\Delta(h,i).$$

Then by (2),
$$D = a - R + (g-1)(a-r) - hR,$$
$$a - 1 - D = (h+1)R - 1 - (g-1)(a-r).$$

The second member is $< a - r$ by (7) and is $\geqq 0$ since

$$(7') \qquad (h+1)R - 1 \geqq (g-1)(a-r),$$

by the definition of g. Since D is the distance of $F(h, Q+g-1)$ from the top of its column and $D \leqq a-1$, the distance of $F(h, Q+g)$ from the top of its column is $D + a - r > a - 1$. Hence $Q + g - 1$ is the largest integer l for which $F(h, l)$ is in the table.

The top F of strip $h+1$ is therefore

$$F(h, Q+g) - a = F(h+1, g).$$

Since $\Delta(h+1, g) = g(a-r) - (h+1)R$ is > 0 by (7) and $< a - r$ by (7′), the induction is complete as to Δ.

Where h is replaced by $h+1$, let g become g', whence g' is the least integer for which

$$(h+2)R - 1 < g'(a-r).$$

Then (7′) gives $g'(a-r) > (g-1)(a-r)$, $g' \geqq g$. By our results about the tops of strips h, $h+1$, we see that i becomes g when h is replaced by $h+1$. Hence $g \geqq i$ follows by induction.

Corollary 1. *If i and g are the least integers for which $F(h, i)$ and $F(h+1, g)$ are in the table, then $g \geqq i$, and $g = i$ or $g = i + 1$.*

To prove the final remark, we add

$$hR \leqq i(a-r) \quad (\text{viz., } \Delta \geqq 0), \qquad R - 1 < a - r$$

and get $(h+1)R - 1 < (i+1)(a-r)$. This with (7′) gives $g - 1 < i + 1$, $g \leqq i + 1$.

Corollary 2. *The number of F's in any strip is $Q - 1$ or Q.*

For, the F's are

$$(9) \qquad F(h, i),\ F(h, i+1), \cdots, F(h, Q+g-1).$$

A useful restatement of part of Theorem 2 is

Corollary 3. *If g is the integer satisfying*

$$(8) \qquad pR - 1 < g(a-r), \qquad pR - 1 \geqq (g-1)(a-r),$$

then $F(p, g)$ is the top F of strip p.

THEOREM 3. *Within any strip h, the weight of any F is less than the weight of any other entry of the same row.*

Let $F(h,s)$ and $F(h,t)$ be any two distinct F's in (9). Let $t < s$. By (6_2),

$$F(h,s) = d + F(h,t), \qquad d = (s-t)(a-r) > 0. \tag{10}$$

The maximum $s-t$ is $\leqq Q$ by Corollary 2. Hence $d \leqq Q(a-r) = a - R$, by (2).

$$\text{Weight } F(h,i) \text{ is } m + k + hQq + iq - h. \tag{11}$$

Thus the weight $d + F(h,t)$ is d plus (11), which sum exceeds the weight $F(h,s)$ if $d > (s-t)q$, viz., $a - r > q$, which is true by Theorem 1.

It remains to treat the entries to the left of $F\ (h,t)$. Transposing d in (10), we see that these are

$$\begin{aligned} a - d &+ F(h,s) - a \\ &= a - d + (m + hQq + hQ + sq + s - h - 1)a + (k - hQ - s)b, \end{aligned}$$

whose weight is $a - d + m + k + hQq + sq - h - 1$. This exceeds the weight of $F(h,t)$ if $a - d + (s-t)q - 1 > 0$. This holds since $a - d \geqq R > 0$ and the remaining part is positive.

THEOREM 4. *Let the weight of the top $F(h,i)$ of strip h exceed $m + k + \Delta$, where $0 \leqq \Delta = i(a-r) - hR < a$. Let C be either the column which contains $F(h,i)$ or any column to the left of it. Then no entry of C is a minimum decomposition.*

Let $F(H,I)$ be in column C. Then $H \geqq h$. If $H = h$, then $I \geqq i$ by the definition of top. If $H > h$, then $I \geqq i$ by Corollary 1. By (5),

$$F(H,I) = \Delta + F(H-h, I-i)$$

is a decomposition. The weight of the first member exceeds that of the second if $hQq + iq - h > \Delta$, which is true by the first hypothesis in Theorem 4. Hence $F(H,I)$ is not a minimum decomposition. The same is evidently true of $j + F(H,I)$ for $0 \leqq j < a$.

The entry at the top of the column containing $F(H,I)$ is the sum of an integer $\geqq 1$ by the function obtained from $F(H,I)$ by subtracting unity from the coefficient of a. Hence that entry is not a minimum decomposition.

By (11) we obtain

LEMMA 1. *The weight of $F(h,i)$ is $>$ or $\leqq m + k + \Delta(h,i)$, according as $E(h,i) > 0$ or $\leqq 0$, where*

(12) $$E(h,i) = hQq + iq - h + hR - i(a-r).$$

LEMMA 2. *The weight of* $F(H,I)$ *is* $\leqq$ *the weight of* $[\Delta(h,i) + F(H-h, I-i)]$ *if and only if* $E(h,i) \leqq 0$.

The value of i for which $F(h,i)$ is the top F of strip h will be denoted by $t(h)$. Define l by

(13) $$E(l+1, t(l+1)) > 0, \quad E(h,t(h)) \leqq 0 \qquad (h = 0, 1, \cdots, l).$$

The strips $l, l-1, \cdots, 1, 0$ will be said to form the *reduced* table. Lemma 1 and Theorem 4 with $h = l+1$ yield

COROLLARY 4. *No entry to the left of the reduced table is a minimum decomposition.*

For $p = 1$, (8) gives $g = 1$, whence $t(1) = 1$. Also $E(0,0) = 0$. Hence $l = 0$ if and only if $E(1,1) > 0$.

COROLLARY 5. *If* $E(1,1) > 0$, *the strip* 0 *forms the reduced table.*

By Theorem 3 and Corollaries 4 and 5, we have

COROLLARY 6. *If* $E(1,1) > 0$, *the only minimum decompositions of integers* $< 4^n$ *are* $c_j + F(0,j)$ *for* $j = 0, \cdots, Q$, *where* $c_j = 0, 1, \cdots, a-r-1$ *if* $j < Q$, *but* $c_Q = 0, \cdots, R-1$.

THEOREM 5. *In the reduced table, the weight of any* F *is* $\leqq$ *the weight of every further entry of the same row.*

By Theorem 3 we may assume that at least two strips occur, and see that it remains only to compare entries in different strips H and $p = H - h$, where $1 \leqq h \leqq H \leqq l$.

Let D_i be the distance of $F_i = F(u, I-i)$ from the top of its column. By (6_2), $F_i = d + F_j$, $d = (j-i)(a-r)$. Since $F_i = D_i + ma + kb$, $D_i - D_j = d$. Hence F_j is above F_i if and only if $j > i$. Thus if an entry is below F_i, it is below F_j for all $j > i$. Hence either (a) $F(H,I)$ is above all F's of strip p, or (b) there is a least i such that $F(H,I)$ is below $F(p, I-i)$ or in the same row with it.

Case (a). $F(H,I)$ is above $F(p,j)$ for each j. If D and D' are their distances from the tops of their columns, then their difference is $D - D' < 0$. Apply (5) with $i = I - j$; thus $F(H,I) - F(p,j) = \Delta(h, I-j)$ and $0 < -\Delta = D' - D < a$. Write d for $-\Delta$. Thus $F(p,j)$ and $F(H,I) + d$ are equal, and the weight of the former will be $<$ the weight of the latter if

$$hQq + (I - j)q - h + d > 0.$$

This evidently holds if $I \geqq j$. Next, let $I - j = -P$, $P > 0$. Inserting the value of d, we obtain the inequality

$$h(Qq - 1 + R) + P(a - q - r) > 0,$$

which follows from Theorem 1.

Consider any integer v smaller than i of Case (b). Then $F(H, I)$ is above $F(p, I - v)$, and the proof in Case (a) evidently applies with $j = I - v$.

Case (b). Write $F(H, I) = D + F(0, 0)$, $F(p, I - i) = D' + F(0, 0)$. Then $D \geqq D'$. By (5),

$$D - D' = F(H, I) - F(p, I - i) = \Delta(h, i), \quad a > \Delta(h, i) \geqq 0.$$

Since $F(H, I)$ is above $F(p, I - i + 1)$, $\Delta(h, i - 1) < 0$. By the inequalities for the two Δ's we see that (8), with p replaced by h, requires $g = i$, so that $F(h, i)$ is the top F of strip h. In other words, $i = t(h)$. But $h \leqq l$. Thus $E(h, i) \leqq 0$ by (13). Thus Lemma 2 shows that the weight of $F(H, I)$ is $\leqq$ the weight of $[\Delta(h, i) + F(p, I - i)]$. This proves the theorem for the present special case.

To this case we shall reduce the proof for an integer $j > i$ such that $F(p, I - j)$ is in our table. Since $F(H, I)$ is below $F(p, I - j)$, we see as at the beginning of Case (b) that $a > \Delta(h, j) > 0$. Denote the weight of $F(p, I - i)$ by w_i. By use of (11), we find that $w_i = w_j + (j - i)q$. We have

$$F(H, I) = F(p, I - j) + \Delta(h, j). \tag{14}$$

The weight of the second member is

$$w_j + \Delta(h, j) = w_i - (j - i)q + \Delta(h, i) + (j - i)(a - r) = w_i + \Delta(h, i) + s,$$

where $s = (j - i)(a - r - q) > 0$ by Theorem 1. By the first paragraph, $\Delta(h, i) + w_i \geqq \text{wt.}\, F(H, I)$. Hence the weight of the second member of (14) exceeds that of the first member.

5. *The typical example* $n = 35$. We find that $13R < 7(a - r)$. If $0 < x \leqq 7$, $7(2x - 1)R \leqq 13xR < 7x(a - r)$. Also $2R > a - r$. Hence

$$(2x - 1)R < x(a - r), \quad (2x - 1)R > (x - 1)(a - r), \qquad 0 < x \leqq 7.$$

Hence for $p = 2x - 1$, (8) require that $g = x$, and Corollary 3 gives $F(2x - 1, x)$. The same inequalities show that $(2x - 2)R$ exceeds $(x - 1)$

$\times (a-r)$ and is $< x(a-r)$, whence $F(2x-2, x)$. Thus for $j=1$ or 2, $F(2x-j, x)$ is the top F of strip $2x-j$. To satisfy (13), we seek the least x for which $E(2x-j, x) > 0$. By (12), the condition is $Ax > B_j$, where

$$A = 5q - 2 + 2R - (a-r) = 1,046,613,414,$$
$$B = 2q - 1 + R = 7,290,593,039.$$

Hence for $j=1$, $x = 6.96$ and the least integer x is 7. For $j=2$, the least integer x is therefore 14. Hence (13) holds for $l+1=13$. Thus the top F's in the reduced table have the subscripts

$$12\,7,\ 11\,6,\ 10\,6,\ 95,\ 85,\ 74,\ 64,\ 53,\ 43,\ 32,\ 22,\ 11,\ 00.$$

6. Condition $E(1,1) > 0$ holds when

$$(15) \qquad n = 2, 3, 5, 6, 8\text{-}12, 14, 15 \qquad (\text{not* for } n = 16\text{-}36).$$

The minima are given by Corollary 6. In the current number of the *Bulletin of the American Mathematical Society*, the minimum weights were compared for the various values of m and k, and complete conclusions drawn as to how many integral n-th powers $\geqq 0$ it is necessary to add together to obtain each integer in various extensive intervals. For each n in (15) the values of $a, q, r, a-r, Q, R$ were tabulated there.

The values of q and r may be found by recursion formulas. Let $3^n = 2^n q(n) + r(n)$, $0 < r(n) < 2^n$. For $q(n) = \text{even} = 2k$, either

$$3r(n) < 2^{n+1}, \quad q(n+1) = 3k, \qquad r(n+1) = 3r(n);$$
$$\text{or} \quad 3r(n) > 2^{n+1}, \quad q(n+1) = 3k+1, \quad r(n+1) = 3r(n) - 2^{n+1}.$$

For $q(n) = \text{odd} = 2k+1$, either

$$2^n + 3r(n) < 2^{n+1}, \quad q(n+1) = q(n) + k, \qquad r(n+1) = 3r(n) + 2^n;$$
$$\text{or} \quad 2^n + 3r(n) > 2^{n+1}, \quad q(n+1) = q(n) + k + 1, \quad r(n+1) = 3r(n) - 2^n.$$

We have $Q=1$ (whence $R=r$) if and only if $a > 2r$. The new cases are

n	q	r
4	5	1
7	17	11
17	985	34 243
20	3325	269 201

n	q	r
24	16 834	1 882 337
25	25 251	5 647 011
27	56 815	17 268 667
29	127 834	21 200 275
30	191 751	63 600 825

$$n = 36, \quad q = 2\,184\,164, \quad r = 28\,111\,390\,417.$$

* Verified direct, but follows also by the sequel.

The new cases with $Q > 1$ are

n	q	r	$a - r$	Q	R
13	194	5 075	3 117	2	1 958
16	656	55 105	10 431	6	2 950
18	1 477	233 801	28 343	9	7 057
19	2 216	439 259	85 029	6	14 114
21	4 987	1 856 179	240 973	8	169 368
22	7 481	3 471 385	722 919	5	579 709
23	11 222	6 219 851	2 168 757	3	1 882 337
26	37 876	50 495 465	16 613 399	4	655 268
28	85 222	186 023 729	82 411 727	3	21 200 275
31	287 626	1 264 544 299	882 939 349	2	381 604 950
32	431 439	3 793 632 897	501 334 399	8	284 292 104
33	647 159	7 085 931 395	1 504 003 197	5	1 069 918 607
34	970 739	12 667 859 593	4 512 009 591	3	3 643 840 411
35	1 456 109	20 823 709 595	13 536 028 773	2	7 287 680 822

7. *The Case* $2R < a - r$. Of importance is the largest integer S satisfying

(16) $$E(S,1) \leqq 0, \quad SR \leqq a - r.$$

Evidently $t(i) = 1$ if $0 < i \leqq S$.

For $n = 16$ or 18, $S = 1$, $E(2,1) > 0$, and $F(1,1), \cdots, F(1,Q)$, $F(0,0), \cdots, F(0,Q)$ are the only F's in the reduced table.

First, let $(S+1)R \leqq a - r$ and (16) hold. By the definition of S as greatest, we have $E(S+1,1) > 0$, whence (13) hold with $l = S$. Hence in the reduced table the only tops F are $F(j,1)$ for $j = 1, \cdots, S$ and $F(0,0)$. The condition holds for $n = 4$, 7, 19, 26; then $S = 2$, 3, 3, 20, respectively.

Second, let $(S+1)R > a - r$. Then $t(S+1) = 2$. We find

n	17	20	24	25	27	28	29	30	31
S	2	2	7	4	6	3	24	15	2
l	10	19	54	58	26	26	96	94	36
L	4	7	7	12	4	7	4	6	16

where $F(l+1, L)$ is the first top to the left of the reduced table. Except for $n = 29$ and $n = 31$, the tops are $F((S+1)x - j, x)$, where $x \geqq 1$ if $j \leqq S$, but $x \geqq 2$ if $j = S + 1$; the details are entirely similar to those in § 5. For $n = 29$, the tops are $F(24(x-1) + j, x)$, with $j = 1, \cdots, 24$ for $x = 1, 2, 3$, but $j = 1, \cdots, 25$ for $x = 4$. For $p = 31$, the tops are $F(2x + p - j, x)$

for $x = 3p + 2$ or $3p + 3$, $j = 0, 1$, and $F(7p + 2 - j, 3p + 1)$, $j = 0, 1, 2$, if $p \geqq 1$, but $j = 0, 1$ if $p = 0$.

8. *The Case* $2R > a - r$. For $n = 13, 32$ or 35, $3R < 2(a - r)$. For $n = 13$, the tops in the reduced table are $F(2, 2)$, $F(1, 1)$ and $F(0, 0)$. For $n = 32$, they are $F(6, 4)$, $F(5, 3)$, $F(4, 3)$, $F(3, 2)$ and the preceding three (laws as in § 5).

For $n = 21$, 33 or 36, $3R > 2(a - r)$, $4R < 3(a - r)$. For $n = 21$, the tops in the reduced table are $F(ii)$, $i = 0, \cdots, 3$. For $n = 36$, the additional tops are $F(12, 9)$, $F(11, 8)$, $F(10, 7)$, $F(9, 7)$, $F(8, 6)$, $F(7, 5)$, $F(6, 5)$, $F(5, 4)$, $F(4, 3)$. For $n = 33$, the tops are

$$F(7x - j, 5x),\quad F(7x + 1, 5x + 1),\quad F(7x + 2, 5x + 2),$$
$$F(7x + 4 - j, 5x + 3),\quad F(7x + 5, 5x + 4)$$

for $j = 0, 1$. The reduced table begins with $F(51, 37)$.

For $n = 22, 23, 34$, $5R > 4(a - r)$. For $n = 22$, the tops in the reduced table are $F(ii)$, $i = 0, \cdots, 5$. For $n = 23$, the tops are $F(ii)$, $i = 0, \cdots, 7$. For $n = 34$, the tops in the reduced table are $F(ii)$, $i = 0, \cdots, 4$ and

$$F(5x, 4x + 1),\ x = 1, \cdots, 4;\qquad F(5x + 1, 4x + 1),\ x = 1, \cdots, 5;$$
$$F(5x - 1 - i, 4x - i),\ i = 0, 1, 2;\ x = 2, 3, 4, 5.$$

The reduced table begins with $F(25, 21)$.

A NOTE ON THE NON-DIFFERENTIABLE FUNCTION OF WEIERSTRASS.

By Aurel Wintner.

The monotone function $\rho(\xi)$, $-\infty < \xi < +\infty$, is said to be the distribution function of the real-valued continuous * function $x(t)$, $-\infty < t < +\infty$, if at every continuity point ξ of ρ

$$\rho(\xi) = \lim_{T=\infty} \{x(t) \leqq \xi;\ T\}/2T$$

where $\{x(t) \leqq \xi;\ T\}$ denotes the measure of the set of those points t at which both inequalities $x(t) \leqq \xi$, $|t| \leqq T$ are satisfied. The existence of a distribution function for any real-valued almost-periodic † function

$$x(t) \sim \sum_{n=1}^{\infty} a_n \cos \lambda_n (t - \phi_n) \tag{1}$$

has originally been proven ‡ in order to make available, in the case of an infinite t-range, an analogue to the measure function of Lebesgue. The latter is obtained by a non-local inversion $t = t(x)$ of $x = x(t)$ and has in the classical case of a finite t-range the object of smoothing the behavior of the original function $x(t)$. In fact, the superiority of the Lebesgue integration theory is due, at least in part, to this inversion. It was, therefore, to be expected that to the rather intricate behavior § of an almost-periodic curve $x = x(t)$ there might correspond an essentially smoother behavior of its distribution function. A proper example has, however, been missing so far.

* This restriction is not a necessary one.

† Almost-periodicity is meant in the original Bohr sense of the word.

‡ A. Wintner, "Diophantische Approximationen und Hermitesche Matrizen. I.," *Mathematische Zeitschrift*, vol. 30 (1929), pp. 310-311. The complex-valued case has then been treated by Jessen under the assumption that the frequencies are linearly independent (Jessen postulates also a restriction which is somewhat stronger than the condition of analyticity but is in reality superfluous in his proof). The complex-valued problem has been solved for arbitrary frequencies and without any analyticity restriction by Haviland. Cf. B. Jessen, *Bidrag til Integral-theorien for Funktioner of uendelig mange Variable*, Copenhagen, 1930, and E. K. Haviland, "On Statistical Methods in the Theory of Almost-periodic Functions," *Proceedings of the National Academy of Sciences*, vol. 19 (1933), pp. 549-555.

§ Cf. O. Toeplitz, "Ein Beispiel zur Theorie der fastperiodischen Funktionen," *Mathematische Annalen*, vol. 98 (1928), p. 281.

From a recent result, the everywhere continuous but nowhere differentiable function of Weierstrass * appears as an illustration of the desired character.†

To see this we need the fact that if all $a_n \neq 0$ and if the frequencies λ_n of (1) are linearly independent then the distribution function of (1) possesses in the whole range $-\infty < \xi < +\infty$ continuous derivatives of arbitrarily high order.‡

The Weierstrass function is

$$(2) \qquad x(t) = \sum_{n=1}^{\infty} a^n \cos b^n t; \qquad 0 < a < 1, \quad 0 < b > C = C_a$$

where the lower bound $C > 0$ of the admissible values of b depends upon a. Since from (1) and (2)

$$\phi_n = 0, \quad a_n = a^n, \quad \lambda_n = b^n,$$

the frequencies λ_n of (1) are linearly dependent if and only if b satisfies a relation $\Sigma\, m_k b^k = 0$ with a finite number of terms where the coefficients m_k are integers and not all zero. Hence on excluding from the admissible range $C_a < b < +\infty$ of b the denumerable set of algebraic numbers, the frequencies of the almost-periodic function (2) will be linearly independent so that *the distribution function of* (2) *everywhere possesses derivatives of arbitrarily high order whereas* (2) *itself is nowhere differentiable and shows a rather intricate behavior not only locally but also in the large.* In fact, the frequencies $\lambda_n = b^n$ are linearly independent so that the curve $x = x(t)$ does not have any intuitive regularity in a large t-range (cf. O. Toeplitz, *loc. cit.*).

The complex-valued function

$$(2a) \qquad x(t) + iy(t) = \sum_{n=1}^{\infty} a^n \exp ib^n t$$

and more generally

$$(1a) \qquad x(t) + iy(t) = \sum_{n=1}^{\infty} a_n \exp i\lambda_n(t - \phi_n)$$

where the frequencies are linearly independent also possesses a distribution function which has, save at most at the origin $x = y = 0$,§ derivatives of arbitrarily high order with respect to x and y.

* Cf. G. H. Hardy, "Weierstrass's non-differentiable function," *Transactions of the American Mathematical Society*, vol. 17 (1916), pp. 301-315. Hardy proves that the function (2) nowhere possesses a finite derivative if $ab \geqq 1$.

† A. Wintner, "Upon a statistical method in the theory of diophantine approximations," *American Journal of Mathematics*, vol. 55 (1933), pp. 309-331.

‡ *Ibid.*, p. 315.

§ *Ibid.*, p. 317. Since then the author has proven that all derivatives exist at the origin also. Cf. (66), p. 325.

The distribution function of (1a) and therefore that of (2a) possesses a *radial symmetry* with respect to the origin.* If one notices this fact, the results regarding the distribution functions of (1) and of (1a) may be shown to be equivalent.† Otherwise ‡ one cannot § deduce *anything* regarding the distribution function of (1) from continuity results regarding the distribution function of (1a).¶

THE JOHNS HOPKINS UNIVERSITY.

---

* *Ibid.*, p. 327. On p. 317 there is cleared up the reason of the apparent paradox pointed out on p. 317 of the author's first paper referred to above.

† *Ibid.*, p. 317.

‡ This is the situation in Jessen's work.

§ Cf. *ibid.*, p. 331. According to a remark of Jessen the author's proof for the statistical independence of the partial distribution functions is essentially the same as the method employed by Bohr in his paper "Another proof of Kronecker's Theorem," *Proceedings of the London Mathematical Society*, ser. 2, vol. 21 (1922), pp. 315-316. Bohr's result is, however, not the so-called Kronecker-Weyl theorem but only the ametrical Kronecker theorem which is unable to yield anything regarding distribution *functions*. It may be mentioned in this connection that the momentum method employed in the proof of the statistical independence yields a direct treatment of the distribution problem of conditionally periodic motions. The usual treatment is based upon Weyl's metrical refinement of the Kronecker theorem.

¶ In his Thesis referred to above, Jessen proves the existence of a continuous mixed derivative $\partial^2/\partial x\partial y = \partial^2/\partial y\partial x$ for the distribution function of (1a). His incidental restriction mentioned above (viz. that (1a) possesses an analytic continuation by means of an analytic almost-periodic function) is not satisfied by the Weierstrass function.

ON THE DISTRIBUTION FUNCTION OF ALMOST-PERIODIC ANGULAR VARIABLES.

By Aurel Wintner.

In the present note the method previously * used in the distribution problem of real-valued almost-periodic functions, i. e. of linear coördinates, will be applied to the corresponding problem regarding angular variables. It will be first shown that every almost-periodic function $f(t)$ for which

$$(1) \qquad |f(t)| = 1, \text{ i. e. } f(t) = \exp\, i\vartheta(t), \quad (-\infty < t < +\infty),$$

possesses a distribution function which will be introduced by means of the trigonometric momentum problem. According to a theorem of Bohr, formulated as a conjecture by the present author, a function $f(t)$ satisfying (1) is almost-periodic if and only if there exist a constant μ and an almost-periodic function $\omega(t)$ such that

$$(2) \qquad \vartheta(t) = \mu t + \omega(t)$$

where μ and $\omega(t)$ are real.† Since the distribution function of $\exp i\omega(t)$ may immediately be obtained from the one which belongs to $\omega(t)$ and since the density of the distribution function of $\exp i\mu t$ is clearly constant, the distribution problem regarding the function

$$(3) \qquad f(t) = \exp i\mu t \cdot \exp i\omega(t)$$

seems to be reducible to the distribution problem of the real-valued almost-periodic function $\omega(t)$ as treated *loc. cit.* Such a reduction is, however, not possible. In fact, there is not known any rule combining the distribution function of the product (3) from the distribution functions of its factors. This situation will be illustrated by an example showing that the distribution function of the second factor in (3) may be discontinuous and the distribution

* A. Wintner, "Diophantische Approximationen und Hermitesche Matrizen. I.," *Mathematische Zeitschrift*, vol. 30 (1929), pp. 290-319.

† H. Bohr, "Kleinere Beiträge zur Theorie der fastperiodischen Funktionen," *Det Kgl. Danske Videnskabernes Selskab. Meddelelser*, vol. 10, no. 10 (1930). The Lagrange-Bohl problem regarding the existence of a mean motion suggests a generalization of the question. Let $g(t)$ be for simplicity the sum of only three vibrations $r_k \exp it(\lambda_k - \phi_k)$. On placing $\exp i\vartheta(t) = g(t)/|g(t)|$ it is known that (2) holds where $\omega(t)$ is $= o(t)$ but not necessarily $= O(1)$, hence not necessarily almost-periodic. Thus it would be interesting to know whether or not $\omega(t)$ is almost-periodic in a generalized sense. Since a real number τ may satisfy both conditions $g(\tau) = 0$, $g'(\tau) = 0$, the function $\vartheta(t)$ may have jumps of modulus π. The ratio $\exp 2i\vartheta(t) = g(t)/\bar{g}(t)$ is, however, regular for real values of t without being necessarily almost-periodic in the original sense of Bohr.

function of the product nevertheless continuous although the distribution function of the other factor in (3) is of constant density. Hence a direct treatment of the functions $f(t)$ satisfying (1) cannot be avoided.*

Let $\sigma(\phi)$, $-\infty < \phi < +\infty$, denote a monotone function satisfying the conditions

$$(4)\quad \sigma(\phi + 2\pi) = \sigma(\phi) + 2\pi,\quad \sigma(\phi - 0) = \sigma(\phi),\quad \sigma(2\pi - 0) = 2\pi.$$

The function σ may be constant in some intervals or it may have discontinuity points so that it need not represent a topological transformation of the circle $|z| = 1$ into itself. On denoting by ψ one of the continuity points of $\sigma(\phi)$, which lie everywhere dense, the value of the Stieltjes integral

$$(5)\qquad \int_{\psi}^{\psi+2\pi} \exp(in\phi)\,d\sigma(\phi) \qquad (n = 0, 1, 2, \cdots)$$

is clearly independent of ψ. The trigonometric momentum problem asks for a solution σ of the infinitely many equations

$$(6)\qquad \int \exp(in\phi)\,d\sigma(\phi) = 2\pi c_n \qquad (n = 0, 1, 2, \cdots)$$

where $c_0 = 1$, $c_1, c_2, \cdots$ is a given sequence of numbers and the integral (6) is an abbreviation for (5). It is known † that this momentum problem possesses a unique monotone solution σ satisfying (4) if and only if the particular Hermite form

$$(7)\qquad \sum_{k=0}^{m} \sum_{l=0}^{m} c_{k-l} x_k \bar{x}_l \quad \text{where} \quad c_{-j} = \bar{c}_j \qquad (j = 0, 1, 2, \cdots)$$

introduced by Toeplitz is non-negative definite for arbitrarily large values of m. It is easy ‡ to see that this condition is satisfied by $c_n = \mathfrak{M}(f^n)$ where f is any almost-periodic function of constant modulus 1 and $\mathfrak{M}$ denotes the time-average operator

$$\mathfrak{M}(\cdots) = \lim_{T=\infty} \int_{-T}^{T} \cdots dt/2T.$$

* The existence of a distribution function for almost-periodic and also for more general classes of angular functions has been proven by the author by means of Cauchy's transform in a paper submitted 1932 to the Monatshefte (not yet appeared). The same result may be deduced from a recent work of Haviland which is based upon analogous considerations. Cf. E. K. Haviland, "On statistical methods in the theory of almost-periodic functions," *Proceedings of the National Academy of Sciences*, vol. 19 (1933), pp. 549-555.

† Cf. G. Herglotz, "Ueber Potenzreihen mit positivem reellen Teil im Einheitskreis," *Sitzungsberichte der Sächsischen Akademie der Wissenschaften zu Leipzig*, vol. 63 (1911), pp. 401-411. Cf. also F. Hausdorff, "Momentenprobleme für ein endliches Integral," *Mathematische Zeitschrift*, vol. 16 (1923), pp. 220-248.

‡ Cf. an analogous application of the Toeplitz forms in the author's paper, "Zur Theorie der beschränkten Bilinearformen," *Mathematische Zeitschrift*, vol. 30 (1929), pp. 228-282.

First, the existence of $\mathfrak{M}(f^n)$ is assured by the fact that f and therefore f^n is almost-periodic. Furthermore, $f(t)^n$ and $f(t)^{-n}$ are conjugated complex inasmuch as $|f(t)|=1$. Hence on placing $c_n=\mathfrak{M}(f^n)$ so that $c_0=1$ in virtue of $\mathfrak{M}(f^0)=\mathfrak{M}(1)=1$, the Toeplitz form (7) may be written as

$$\sum_{k=0}^{m}\sum_{l=0}^{m}\mathfrak{M}(f^{k-l})x_k\bar{x}_l=\mathfrak{M}(\sum_{k=0}^{m}\sum_{l=0}^{m}f^{k-l}x_k\bar{x}_l)$$
$$=\mathfrak{M}(\sum_{k=0}^{m}\sum_{l=0}^{m}f^k\bar{f}^l x_k\bar{x}_l)=\mathfrak{M}(|\sum_{k=0}^{m}f^k x_k|^2)$$

and is therefore everywhere $\geqq 0$; q. e. d. Thus there exists for every almost-periodic function $f(t)$ of constant modulus 1 exactly one monotone function $\sigma(\phi)$ satisfying (4) such that

$$(8)\qquad \int \exp(in\phi)\,d\sigma(\phi)=2\pi\mathfrak{M}(f^n)\qquad (n=0,1,2,\cdots).$$

This function $\sigma(\phi)$ will be termed the distribution function belonging to $f(t)$. For $n=0$ we have

$$(9)\qquad \int d\sigma(\phi)=2\pi\mathfrak{M}(f^0)=2\pi=\sigma(2\pi)$$

in virtue of (8) and (4).

We shall now justify the name "distribution function."

Let $\{\vartheta(t)\}$ denote the least non-negative remainder mod 2π of the continuous arcus (2) of $f(t)$. Let ϕ be any positive number $<2\pi$ and let $\{\phi,T\}$ denote the measure of the set of those points in the range $-T\leqq t\leqq T$ at which $0\leqq\{\vartheta(t)\}\leqq\phi$. The function $\chi(\phi;\,T)$ defined for

$$(10)\qquad 0\leqq\phi<2\pi$$

as

$$(11)\qquad \chi(0;\,T)=0;\qquad \chi(\phi;\,T)=\{\phi;\,T\}/2T,\quad 0<\phi<2\pi$$

is monotone in the range (10) and satisfies the relations *

$$(12)\qquad \int_0^{2\pi-0}\exp(in\phi)\,d\chi(\phi;\,T)=2\pi\int_{-T}^{T}f(t)^n\,dt/2T\qquad(n=0,1,2,\cdots)$$

inasmuch as the Stieltjes approximative sums of the first integral are, up to the factor $2\pi/2T$, precisely the Lebesgue approximative sums of the second integral. On defining $\chi(\phi;\,T)$ outside of the range (10) by means of the relation

$$(13)\quad \chi(\phi+2\pi;\,T)=\chi(\phi;\,T)+2k\pi\quad(0\leqq\phi<2\pi;\ k=\pm1,\pm2,\cdots)$$

we have $\chi(2\pi;\,T)=2\pi$ inasmuch as $\chi(0;\,T)=0$ in virtue of (11). On the other hand, (12) yields for $n=0$ that

* The Stieltjes integration concerns ϕ whereas T has a fixed value.

(13a) $$\chi(2\pi-0;\ T)-\chi(0;\ T)=\chi(2\pi-0;\ T)=2\pi.$$

Hence in (12)

(14) $$\int_0^{2\pi-0}=\int_0^{2\pi}.$$

We are now in a position to prove * that there exists a monotone function $\rho(\phi)$, $-\infty<\phi<+\infty$ such that

(15) $$\lim_{T=\infty}\chi(\phi;\ T)=\rho(\phi)$$

at all continuity points ϕ of ρ. Suppose, if possible, the contrary. According to the compactness theorem of Helly † there will then exist two monotone non-bounded sequences $\{T'_k\}$, $\{T_k''\}$ such that the limits

(16) $$\rho_1(\phi)=\lim_{k=\infty}\chi(\phi;\ T'_k),\qquad \rho_2(\phi)=\lim_{k=\infty}\chi(\phi;\ T_k'')$$

exist and represent two monotone functions which are such that $\rho_1(\phi)=\rho_2(\phi)$ does not hold at every continuity point of ρ_1 or ρ_2. It is clear from the definition of $\chi(\phi;\ T)$ that we may confine ourselves to the finite range (10). From (12), (14) and (16) we have for every n and for $\nu=1,2$

$$\int_0^{2\pi}\exp(in\phi)\,d\rho_\nu(\phi)=2\pi\mathfrak{M}(f^n)$$

in virtue of the Helly ‡ theorem on term-by-term integration. It follows therefore from the uniqueness theorem regarding the solution of the trigonometric momentum problem that the difference $\rho_1(\phi-0)-\rho_2(\phi-0)$ is a constant. This constant is, however, equal to zero, inasmuch as $\rho_1(2\pi-0)=\rho_2(2\pi-0)$ in virtue of $2\pi\mathfrak{M}(f^0)=2\pi$. Consequently ρ_1 and ρ_2 are identical at all their continuity points. Since this is a contradiction, the statement (15) is proven.

From (4), (13a), (13) and (15) we have

$$\sigma(2\pi-0)=\rho(2\pi-0)\quad\text{and}\quad \rho(\phi+2\pi)=\rho(\phi)+2\pi.$$

Furthermore, from (12), (14) and (15)

$$\int_0^{2\pi}\exp(in\phi)\,d\rho(\phi)=2\pi\mathfrak{M}(f^n)\qquad(n=0,1,2,\cdots)$$

in virtue of the Helly theorem on term-by-term integration. Since (8) cannot have more than one monotone solution σ satisfying (4), it follows that $\rho(\phi)=\sigma(\phi)$ at all continuity points of σ. Hence from (15), (11), (4) and (13)

---

* Cf. p. 105 of the author's book "Spektraltheorie der unendlichen Matrizen," Leipzig, 1929.

† E. Helly, "Ueber lineare Funktionaloperationen," *Sitzungsberichte der mathematisch-naturwissenschaftlichen Klasse der Kaiserlichen Akademie der Wissenschaften zu Wien*, vol. 121 (1912), pp. 265-297.

‡ E. Helly, *loc. cit.*

$$\sigma(\phi) = \lim_{T=\infty} \chi(\phi;\, T), \tag{17}$$

at least if ϕ is neither a point of the at most enumerable set of the discontinuity points of σ nor * of the form $2\pi k$ $(k = 0, \pm 1, \pm 2, \cdots)$.

It follows from (17) and from the definition of the measure $\{\phi;\, T\}$ that the function σ, originally defined by means of (8), describes the asymptotic repartition of the values taken by $f(t)$ when $t \to \infty$. The name "distribution functions" is therefore now justified. It follows by a simple modification of an example constructed by Bohr † that (17) need not hold at a discontinuity point of σ. It is, however, possible that (17) holds even at a discontinuity point $\phi \neq 2\pi k$ of σ. This is e. g. the case at $\phi = \pi$ for the periodic function $f(t) = -\exp i\omega(t)$ where

$$\omega(t) = (t^2 - 1)^2 \text{ when } |\, t \,| \leqq 1, \qquad \omega(t) = 0 \text{ when } 1 \leqq |\, t \,| \leqq \pi,$$
$$\omega(t + 2\pi k) = \omega(t) \text{ when } -\infty < t < +\infty.$$

In fact, $\exp i\omega(t)$ is equal to 1 in the periodic images of the range $\pi - 1 \leqq t \leqq \pi + 1$ so that the distribution function of $\exp i\omega(t)$ is discontinuous at $\phi = 0$. On the other hand, the periodic function

$$\exp i\vartheta(t) = \exp it \cdot \exp i\omega(t) \quad \text{where} \quad \vartheta(t) = \mu t + \omega(t) \quad \text{and} \quad \mu = 1$$

is nowhere constant and has an everywhere continuous distribution function. Hence it is possible that the distribution function of the product (3) be everywhere continuous whereas the second factor of the product (3) has a discontinuous distribution function. In other words, the secular term μt of uniform angular distribution is able to dissolve a discontinuity of the original distribution.

It would be desirable to extend, by means of Radon integrals,‡ the proof for the existence of an angular distribution function to the case where the curve $f = f(t)$ lies not on the one-dimensional manifold $|\, z \,| = 1$ but in a more general way on the n-dimensional torus resulting from the n-dimensional euclidean space by reduction mod 1. Such an extension would yield a generalization of the Kronecker-Weyl approximation theorem to cases where the asymptotic distribution is not a uniform one. The same extension is of interest also in connection with the Poincaré differential equation $dy/dx = g(x, y)$ where g is doubly periodic $(n = 2)$.

The Johns Hopkins University.

---

* The function $\chi(\phi;\, T)$ was defined at $\phi = 2\pi k$ not $= \{\phi;\, T\}/2T$ but $= 2\pi k$; cf. (11) and (13).

† H. Bohr, *loc. cit.*

‡ Cf. E. K. Haviland, *loc. cit.*

CONCERNING PRIMITIVE GROUPS OF CLASS U; PAPER II.

By C. F. Luther.

In the preceding paper * limits to the degree n of multiply transitive groups of class u containing a substitution of order 2 and degree $u+\epsilon$ (ϵ a positive integer) are given. The development of these limits depends upon an auxiliary theorem concerning the maximum degree of diedral rotation groups of class u generated by two substitutions s and t of order 2 and degree $u+\epsilon$. Three cases arise: first, the order of st is an odd number; second, the order of st is twice an odd number; third, the order of st is divisible by 4. The second case is found to give the most unfavorable limit; and, when applied to multiply transitive groups, is largely responsible for the undue prominence of the ϵ terms in the limits obtained.

Professor W. A. Manning suggested that if higher transitivity were to be used and some sacrifice be made in the coefficient of u, the ϵ term could be diminished; for higher transitivity would permit dependence upon the third part of the auxiliary theorem alone, and that part, it was hoped, in an improved form. These suggestions are considered in this paper and the results are gratifying. The third case of the auxiliary theorem is now covered by a stronger theorem and decided improvements, both in form and actual value, upon the limits of the preceding paper are at once apparent. Before, in the general case, the coefficient of ϵ was greater than 2 and might increase indefinitely with increasing multiplicity of transitivity; now, it has a maximum of 2 with an asymptotic value of 1.

The principal results are:

Theorem I. *If n is the degree of a more than 2^α ($\alpha \geqq 2$) times transitive group, not alternating or symmetric, that contains a substitution of degree v and order 2, then*

$$n \leqq 2^\alpha v/(2^\alpha - 2).$$

Theorem II. *If n is the degree and u (>3) is the class of a more than $2^\alpha + p_1 + p_2 + \cdots + p_r$ times transitive group ($\alpha \geqq 2$, $p_1, p_2, \cdots, p_r$ dis-*

* Luther, *American Journal of Mathematics*, Vol. 55 (1933), pp. 77-101.

tinct odd primes, $r \geqq 1$) that contains a substitution of degree $u + \epsilon$ and order 2,

$$n < \frac{2^a p_1 p_2 \cdots p_r}{2^a p_1 p_2 \cdots p_r - 2} u + \frac{2^a}{2^a - 2} \epsilon + 1.$$

For the proof of these two theorems it is first necessary to prove:

THEOREM III. *If s and t are two substitutions of order 2 and degree $u + \epsilon$ that generate a group $\{s, t\}$ of degree n and class u, and if the order of st is divisible by $2^a (\alpha \geqq 2)$ and by each of the odd prime power factors $p_1^{a_1}, p_2^{a_2}, \cdots, p_r^{a_r}$ $(r \geqq 0)$,*

$$n \leqq \frac{2^{a-1} p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r} + 1}{2^{a-1} p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}} u + \frac{2^{a-1} + 1}{2^{a-1}} \epsilon.$$

Use is made of a method devised by Professor Manning * for the case $\epsilon = 0$ and used by the author † in the paper to which this is a sequel, for the general case of $\epsilon > 0$. Formulas from the preceding paper will be used here whenever applicable.

Consider the group generated by s and t. Professor Manning has shown that the deletion of all regular constituents on letters common to s and t does not affect the truth of our theorem. Therefore, in what follows we assume the group free of all such regular constituents.

Let st be of order $2^a (\alpha \geqq 2)$. Let $\{s, t\}$ have m_1 transitive constituents of degree 2, m_2 transitive constituents of degree $2^2, \cdots, m_a$ transitive constituents of degree 2^a. From st and $(st)^{2^{a-1}}$ we have:

$$\sum_{i=1}^{a} 2^i m_i = u + \delta$$
$$2^a m_a = u + H.$$

Formula (3) (preceding paper) gives the third equation:

$$\sum_{i=1}^{a} (2^i - 1) m_i = u + \epsilon.$$

Eliminate m_1 and m_a from these three equations:

---

* Manning, *Transactions of the American Mathematical Society*, Vol. 18 (1917), pp. 464 ff.

† Luther, *American Journal of Mathematics*, Vol. 55 (1933), pp. 78-80.

$$\begin{vmatrix} 2 & 2^a & u+\delta-\sum_{i=2}^{a-1} 2^i m_i \\ 0 & 2^a & u+H \\ 1 & 2^a-1 & u+\epsilon-\sum_{i=2}^{a-1}(2^i-1)m_i \end{vmatrix} = 0.$$

Expanding this:

$$\delta/2 = u/2^a + \epsilon - H/2 + H/2^a + \sum_{i=2}^{a-1} m_i - \sum_{i=2}^{a-1} 2^{i-1} m_i.$$

But if we assume, as we legitimately may, that there exists no substitution of order 2 of degree less than $u+\epsilon$, then $H \geqq \epsilon$. Hence

$$\delta \leqq 2^{1-a}u + \epsilon + 2^{1-a}\epsilon - \sum_{i=2}^{a-1}(2^i-2)m_i;$$
$$\leqq 2^{1-a}u + (1+2^{1-a})\epsilon.$$

Then

$$n \leqq (1+2^{1-a})(u+\epsilon).$$

In the general case the order of st is $2^a\pi$, where $\pi = p_1^{a_1}p_2^{a_2}\cdots p_r^{a_r}$, the product of $r(\geqq 1)$ odd prime power factors, and as before $a \geqq 2$. Let there be m_i transitive constituents of degree 2^i, $i=1,\cdots,a$; and $y^{(h)}_{ij\ldots t}$ transitive constituents of degree

$$2^h p_1^i p_2^j \cdots p_r^t;\ h=0,1,\cdots,a;\ i=0,\cdots,a_1;$$
$$j=0,\cdots,a_2;\ \cdots;\ t=0,\cdots,a_r.$$

In the second terms of I, II, and IV below, $i, j, \cdots, t$ are not all zero at the same time. st and $(st)^{2^{a-1}\pi}$ give:

$$\text{I.}\quad \sum_{i=1}^{a} 2^i m_i + \sum_{h,i,j,\ldots,t=0}^{a,a_1,a_2,\ldots,a_r} 2^h p_1^i p_2^j \cdots p_r^t y^{(h)}_{ij\ldots t} = u+\delta$$

$$\text{II.}\quad 2^a m_a + \sum_{i,j,\ldots,t=0}^{a_1,a_2,\ldots,a_r} 2^a p_1^i p_2^j \cdots p_r^t y^{(a)}_{ij\ldots t} = u+H.$$

For a third equation take $1/r$ times the sum of the r equations giving the degrees of the r substitutions $(st)^{2^a\pi/p_1},\cdots,(st)^{2^a\pi/p_r}$, namely:

$$\text{III.}\quad (1/r)\sum_{v=1}^{r} p_v^{av} \sum_{h,i,j,\ldots,t=0}^{a,a_1,a_2,\ldots,a_r} 2^h p_1^i p_2^j \cdots p^m{}_{v-1} p^n{}_{v+1} \cdots p_r^t y^{(h)}_{ij\ldots ma_vn\ldots t} = u+K.$$

For a fourth equation use $\Gamma_1 - \Gamma_0 = \epsilon - \delta + m$ (formulas (3) and (4), preceding paper), which now is

$$-\sum_{i=1}^{a} m_i - \sum_{h,i,\ldots,t=0}^{a,a_1,\ldots,a_r} y^{(h)}_{ij\ldots t} = \epsilon - \delta.$$

Subtract II from I; follow by subtracting III from II; the elimination

of m_1, m_a, and $y^{(a)}_{a_1\ldots a_r}$ from the four equations gives:

$$\left|\begin{array}{llll} 2 & 0 & 0 & \delta - H - \sum_{i=2}^{a-1} 2^i m_i - \sum_{h,i,j,\ldots,t=0}^{a,a_1,a_2,\ldots,a_r} 2^h p_1{}^i p_2{}^j\cdots p_r{}^t y^{(h)}_{ij\ldots t} \\ 0 & 2^a & 0 & H - K - \sum'^{\,a_1,a_2,\ldots,a_r}_{i,j,\ldots,t=0} 2^a p_1{}^i p_2{}^j\cdots p_r{}^t y^{(a)}_{ij\ldots t} \\ & & & \quad + (1/r)\sum_{v=1}^{r} p_v{}^{a_v} \\ & & & \quad \times \sum'^{\,a,a_1,\ldots,a_r}_{h,i,\ldots,t=0} 2^h p_1{}^i\cdots p^m{}_{v-1}p^n{}_{v+1}\cdots p_r{}^t y^{(h)}_{i\ldots m a_v n\ldots t} \\ 0 & 0 & 2^a\pi & u + K - (1/r)\sum_{v=1}^{r} p_v{}^{a_v} \\ & & & \quad \times \sum'^{\,a,a_1,\ldots,a_r}_{h,i,\ldots,t=0} 2^h p_1{}^i\cdots p^m{}_{v-1}p^n{}_{v+1}\cdots p_r{}^t y^{(h)}_{i\ldots m a_v n\ldots t} \\ -1 & -1 & -1 & \epsilon - \delta + \sum_{i=2}^{a-1} m_i + \sum'^{\,a,a_1,\ldots,a_r}_{h,i,\ldots,t=0} y^{(h)}_{i\ldots t} \end{array}\right| = 0.$$

The Σ' indicates that the term containing $y^{(a)}_{a_1a_2\ldots a_r}$ is missing from the sum. Expanding:

$$\delta/2 = u/2^a\pi + \epsilon - H/2 + H/2^a$$
$$- K/2^a + K/2^a\pi - \sum_{i=2}^{a-1}(2^{i-1}-1)m_i - \Phi - \Phi'.$$

The terms Φ and Φ' are given below:

$$\Phi = \sum_{h=0}^{a-1}\Bigg\{\sum_{i,j,\ldots,t=0}^{a_1,a_2,\ldots,a_r}(2^{h-1}p_1{}^i p_2{}^j\cdots p_r{}^t - 1)y^{(h)}_{ij\ldots t} - (1-1/\pi)(2^h/2^a r)$$
$$\times\Bigg[\sum_{j,\ldots,t=0}^{a_2,\ldots,a_r} p_1{}^{a_1}p_2{}^j\cdots p_r{}^t y^{(h)}_{a_1 j\ldots t} + \sum_{i,k,\ldots,t=0}^{a_1,a_3,\ldots,a_r} p_1{}^i p_2{}^{a_2} p_3{}^k\cdots p_r{}^t y^{(h)}_{i a_2 k\ldots t}$$
$$+\cdots+\sum_{i,j,\ldots,s=0}^{a_1,a_2,\ldots,a_{r-1}} p_1{}^i p_2{}^j\cdots p^s{}_{r-1}p_r{}^{a_r} y^{(h)}_{ij\ldots s a_r}\Bigg]\Bigg\}$$
$$\geqq \sum_{h=0}^{a-1}\Bigg\{\Bigg[\sum_{j,\ldots,t=0}^{a_2,\ldots,a_r}(2^{h-1}p_1{}^{a_1}p_2{}^j\cdots p_r{}^t - 1)y^{(h)}_{a_1 j\ldots t}$$
$$+\cdots+\sum_{i,j,\ldots,s=0}^{a_1,a_2,\ldots,a_{r-1}}(2^{h-1}p_1{}^i p_2{}^j\cdots p^s{}_{r-1}p_r{}^{a_r} - 1)y^{(h)}_{ij\ldots s a_r}\Bigg]$$
$$-(1-1/\pi)(2^h/4r)\Bigg[\sum_{j,\ldots,t=0}^{a_2,\ldots,a_r} p_1{}^{a_1}p_2{}^j\cdots p_r{}^t y^{(h)}_{a_1 j\ldots t}$$
$$+\cdots+\sum_{i,j,\ldots,s=0}^{a_1,\ldots,a_{r-1}} p_1{}^i p_2{}^j\cdots p^s{}_{r-1}p_r{}^{a_r} y^{(h)}_{ij\ldots s a_r}\Bigg]\Bigg\}$$
$$\geqq \sum_{h=0}^{a-1}\Bigg\{p_1{}^{a_1}\sum_{j,\ldots,t=0}^{a_2,\ldots,a_r}[2^{h-1}(1-1/2r+1/2\pi r)p_2{}^j\cdots p_r{}^t - 1/p_1{}^{a_1}]y^{(h)}_{a_1 j\ldots t}$$
$$+\cdots+p_r{}^{a_r}\sum_{i,\ldots,s=0}^{a_1,\ldots,a_{r-1}}[2^{h-1}(1-1/2r+1/2\pi r)p_1{}^i\cdots p^s{}_{r-1}p_r{}^{a_r}$$
$$-1/p_r{}^{a_r}]y^{(h)}_{i\ldots s a_r}\Bigg\}$$

$$\geqq \sum_{h=0}^{a-1} \Big\{ \sum_{j,\ldots,t=0}^{a_2,\ldots,a_r} [2^{h-1}(1-1/2r+1/2\pi r)-1/p_1^{a_1}] y^{(h)}_{a_1 j\ldots,t}$$

$$+\cdots+ \sum_{i,\ldots,s=0}^{a_1,\ldots,a_{r-1}} [2^{h-1}(1-1/2r+1/2\pi r)-1/p_r^{a_r}] y^{(h)}_{i\ldots s a_r} \Big\}$$

$$\geqq \sum_{h=0}^{a-1} 2^h \Big\{ (1/2-1/4r+1/4\pi r-1/2^h p_1^{a_1}) \sum_{j,\ldots,t=0}^{a_2,\ldots,a_r} y^{(h)}_{a_1 j\ldots t}$$

$$+\cdots+ (1/2-1/4r+1/4\pi r-1/2^h p_r^{a_r}) \sum_{i,\ldots,s=0}^{a_1,\ldots,a_{r-1}} y^{(h)}_{i\ldots s a_r} \Big\}$$

$$\geqq \sum_{h=0}^{a-1} \Big\{ (1/2+1/4\pi r-1/4r-1/p_1^{a_1}) \sum_{j,\ldots,t=0}^{a_2,\ldots,a_r} y^{(h)}_{a_1 j\ldots t}$$

$$+\cdots+ (1/2+1/4\pi r-1/4r-1/p_r^{a_r}) \sum_{i,\ldots,s=0}^{a_1,\ldots,a_{r-1}} y^{(h)}_{i\ldots s a_r} \Big\}$$

$$\geqq 0.$$

The last term, Φ', is

$$\sum_{i=1}^{a_1-1} (p_1^{i}-1) y_i^{(a)} \text{ when } r=1. \quad \text{When } r \geqq 2,$$

$$\Phi' = \sum_{i,\ldots,t=0}^{a_1,\ldots,a_r}{}' (p_1^{i} p_2^{j} \cdots p_r^{t}-1) y^{(a)}_{ij\ldots t} - (1/r)(1-1/\pi)$$

$$\times \Big[\sum_{j,\ldots,t=0}^{a_2,\ldots,a_r}{}' p_1^{a_1} p_2^{j} \cdots p_r^{t} y^{(a)}_{a_1 j\ldots t} + \cdots + \sum_{i,\ldots,s=0}^{a_1,\ldots,a_{r-1}}{}' p_1^{i} \cdots p^{s}_{r-1} p_r^{a_r} y^{(a)}_{i\ldots s a_r} \Big]$$

$$\geqq \sum_{j,\ldots,t=0}^{a_2,\ldots,a_r}{}' (p_1^{a_1} p_2^{j} \cdots p_r^{t} - 1 - p_1^{a_1} p_2^{j} \cdots p_r^{t}/r) y^{(a)}_{a_1 j\ldots t}$$

$$+\cdots+ \sum_{i,\ldots,s=0}^{a_1,\ldots,a_{r-1}}{}' (p_1^{i} \cdots p^{s}_{r-1} p_r^{a_r} - 1 - p_1^{i} \cdots p^{s}_{r-1} p_r^{a_r}/r) y^{(a)}_{i\ldots s a_r}$$

$$\geqq \sum_{j,\ldots,t=0}^{a_2,\ldots,a_r}{}' [(1-1/r) p_1^{a_1} - 1] y^{(a)}_{a_1 j\ldots t}$$

$$+\cdots+ \sum_{i,\ldots,s=0}^{a_1,\ldots,a_{r-1}}{}' [(1-1/r) p_r^{a_r} - 1] y^{(a)}_{i\ldots s a_r}$$

$$\geqq 0.$$

As before, $H \geqq \epsilon$, and $K \geqq 0$. Therefore $\delta \leqq u/2^{a-1}\pi + \epsilon + \epsilon/2^{a-1}$; and finally

$$n \leqq u + \epsilon + u/2^{a-1}\pi + \epsilon/2^{a-1}.$$

Proof of Theorem I.

Let G, a group of degree n and class $u(>3)$, be more than 2^a times transitive, $\alpha \geqq 2$. By hypothesis there is in it a substitution s of order 2 and degree $v=(u+\epsilon)$:

$$s = (a_1 a_2)(a_3 a_4) \cdots (a_{2^a-1} a_{2^a}) \cdots$$

and a similar substitution,

$$s' = (a_1 a_3)(a_2 a_5) \cdots (a_{2^a-4} a_{2^a-1}) \cdots (a_{2^a})(a_{2^a-2})$$

such that ss' is of order 2^a or a multiple of 2^a. Since G is more than 2^a

times transitive, it contains a transitive subgroup G_{2a} fixing the 2^a letters $a_1, a_2, \cdots, a_{2^a}$. Transforming s' by G_{2^a} gives a set of g_{2^a} substitutions $s', s'', \cdots, s^{(g_{2^a})}$ such that every product $ss^{(i)}$ contains a cycle of order 2^a. We make use of Theorem III in the following way:

$$x_i + 2^a - 2 = \text{number of letters common to } s \text{ and } s^{(i)}.$$

Let
$$z_i = \text{number of letters of } s^{(i)} \text{ new to } s.$$

Hence

$$z_i \leqq v/2^{a-1},$$

from which it follows that

$$x_i \geqq (v - 2^a)(1 - 2^{1-a}).$$

Hence

$$\sum_{g_{2^a}} x_i \geqq (v - 2^a)(1 - 2^{1-a}) g_{2^a}.$$

Now any one of the last $v - 2^a$ letters of s is found in exactly

$$(v - 2^a + 2) g_{2^a} / (n - 2^a)$$

of the substitutions $s', s'', \cdots, s^{(g_{2^a})}$.

Then

$$\sum_{g_{2^a}} x_i = (v - 2^a)(v - 2^a + 2) g_{2^a} / (n - 2^a).$$

Hence

$$(v - 2^a)(v - 2^a + 2)/(n - 2^a) \geqq (v - 2^a)(1 - 2^{1-a}),$$

from which
$$n \leqq 2^a v / (2^a - 2).$$

It may be of interest to note what this limit becomes when the transitivity, t, is introduced. By hypothesis $2^a < t \leqq 2^{a+1}$ $(\alpha \geqq 2)$. Hence

$$n \leqq tv/(t - 4).$$

Further, G contains a doubly transitive subgroup of degree $n - t + 2$, and in it there is a substitution of degree $n - t + 2$. Therefore,

$$n \leqq t(n - t + 2)/(t - 4).$$

Hence

$$n \geqq (t^2 - 2t)/4 \qquad (t > 4).$$

While this formula is inferior to that of Professor Marie Weiss,* it may be useful because of its simplicity.

Proof of Theorem II.

Let G be more than $2^a + \sigma$ times transitive, where

---

* M. J. Weiss, *Transactions of the American Mathematical Society*, Vol. 32 (1930), pp. 262-263.

$$\sigma = p_1 + p_2 + \cdots + p_r,$$

the sum of $r(\geqq 1)$ odd primes, and as before $\alpha \geqq 2$. There exists in G a substitution of order 2 and degree $u + \epsilon$:

$$s = (a_1a_2)(a_3a_4) \cdots (a_{2^\alpha-1}a_{2^\alpha})(b_1b_2) \cdots (b_{p_1-2}b_{p_1-1}) \cdots (d_1d_2) \cdots (d_{p_r-2}d_{p_r-1}) \cdots (b_{p_1}) \cdots (d_{p_r}) \cdots,$$

and a second substitution,

$$s' = (a_1a_3)(a_2a_5) \cdots (a_{2^\alpha-4}a_{2^\alpha-1})(b_1b_3) \cdots (b_{p_1-3}b_{p_1}) \cdots (d_1d_3) \cdots (d_{p_r-3}d_{p_r}) \cdots (a_{2^\alpha})(a_{2^\alpha-2})(b_{p_1-1}) \cdots (d_{p_r-1}) \cdots.$$

For simplicity replace $2^\alpha + \sigma$ by T. Transform s' by G_T to give $s', s'', \cdots, s^{(g_T)}$. Now,

$$x_i + T - 2r - 2 = \text{number of letters common to } s \text{ and } s^{(i)}.$$
$$z_i = \text{number of letters of } s^{(i)} \text{ new to } s.$$

By Theorem III,

$$x_i \geqq u + \epsilon - u/2^{\alpha-1}\prod_1^r p_k - \epsilon/2^{\alpha-1} - T + 2r + 2.$$

So,

$$\sum_{g_T} x_i \geqq \{(2^{\alpha-1}\prod_1^r p_k - 1)u/2^{\alpha-1}\prod_1^r p_k + (2^{\alpha-1} - 1)\epsilon/2^{\alpha-1} - T + 2r + 2\}g_T.$$

As before, any one of the last $u + \epsilon - T + r$ letters of s is found in $(u + \epsilon - T + r + 2)g_T/(n - T)$ of the substitutions $s', s'', \cdots, s^{(g_T)}$. Therefore,

$$\sum_{g_T} x_i = (u + \epsilon - T + r)(u + \epsilon - T + r + 2)g_T/(n - T),$$

or,

$$(u + \epsilon - T + r)(u + \epsilon - T + r + 2)/(n - T) \geqq (2^{\alpha-1}\Pi - 1)u/2^{\alpha-1}\Pi + (2^{\alpha-1} - 1)\epsilon/2^{\alpha-1} - T + 2r + 2,$$

where

$$\Pi = \prod_1^r p_k.$$

If $n = 2^{\alpha-1}\Pi\, u/(2^{\alpha-1}\Pi - 1) + 2^{\alpha-1}\epsilon/(2^{\alpha-1} - 1) + 1$ fails to satisfy the preceding inequality, we can say that,

$$n < 2^{\alpha-1}\Pi\, u/(2^{\alpha-1}\Pi - 1) + 2^{\alpha-1}\epsilon/(2^{\alpha-1} - 1) + 1.$$

We proceed to show that,

$$[(2^{\alpha-1}\Pi\, u/(2^{\alpha-1}\Pi - 1) + 2^{\alpha-1}\epsilon/(2^{\alpha-1} - 1) - T + 1] \times [(2^{\alpha-1}\Pi - 1)u/2^{\alpha-1}\Pi + (2^{\alpha-1} - 1)\epsilon/2^{\alpha-1} - T + 2r + 2] > (u + \epsilon - T + r)(u + \epsilon - T + r + 2).$$

Expanding,

$$\{(2^{\alpha-1}-1)\Pi/(2^{\alpha-1}\Pi-1)+(2^{\alpha-1}\Pi-1)/(2^{\alpha-1}-1)\Pi\}u\epsilon$$
$$+\{(2r+2)2^{\alpha-1}\Pi/(2^{\alpha-1}\Pi-1)+(2^{\alpha-1}\Pi-1)/2^{\alpha-1}\Pi$$
$$-T[2^{\alpha-1}\Pi/(2^{\alpha-1}\Pi-1)+(2^{\alpha-1}\Pi-1)/2^{\alpha-1}\Pi]\}u$$
$$+\{(2r+2)2^{\alpha-1}/(2^{\alpha-1}-1)+(2^{\alpha-1}-1)/2^{\alpha-1}$$
$$-T[2^{\alpha-1}/(2^{\alpha-1}-1)+(2^{\alpha-1}-1)/2^{\alpha-1}]\}\epsilon-T+2$$
$$>r^2+2ru+2u+2r\epsilon+2\epsilon+2u\epsilon-2uT-2\epsilon T.$$

Now,

$$(2^{\alpha-1}-1)\Pi/(2^{\alpha-1}\Pi-1)+(2^{\alpha-1}\Pi-1)/(2^{\alpha-1}-1)\Pi$$
$$=2+(\Pi-1)^2/\Pi(2^{\alpha-1}\Pi-1)(2^{\alpha-1}-1)$$

and also,

$$2^{\alpha-1}/(2^{\alpha-1}-1)+(2^{\alpha-1}-1)/2^{\alpha-1}=2+1/2^{\alpha-1}(2^{\alpha-1}-1),$$

so that the above becomes,

$$\{(\Pi-1)^2/\Pi(2^{\alpha-1}\Pi-1)(2^{\alpha-1}-1)\}u\epsilon$$
$$+\{(2r+2)/(2^{\alpha-1}\Pi-1)+(2^{\alpha-1}\Pi-1)/2^{\alpha-1}\Pi-T/2^{\alpha-1}\Pi(2^{\alpha-1}\Pi-1)\}u$$
$$+\{(2r+2)/(2^{\alpha-1}-1)+(2^{\alpha-1}-1)/2^{\alpha-1}-T/2^{\alpha-1}(2^{\alpha-1}-1)\}\epsilon-T+2>r^2.$$

It is known that the class of a t-ply transitive (non-alternating) group cannot be less than $2t-2$. Hence, we can say that $u\geqq 2T$ in the above inequality. Then,

$$\{2(\Pi-1)^2T/\Pi(2^{\alpha-1}\Pi-1)(2^{\alpha-1}-1)$$
$$+(2r+2)/(2^{\alpha-1}-1)-T/2^{\alpha-1}(2^{\alpha-1}-1)+(2^{\alpha-1}-1)/2^{\alpha-1}\}\epsilon$$
$$+\{(2r+2)/(2^{\alpha-1}\Pi-1)+1-1/2^{\alpha-1}\Pi$$
$$-T/2^{\alpha-1}\Pi(2^{\alpha-1}\Pi-1)-1/2\}2T>r^2-2.$$

The coefficient of ϵ is positive, because

$$2(\Pi-1)^2T/2^{\alpha-1}\Pi^2+2r+2>T/2^{\alpha-1},$$

for $$T-4T/\Pi+2^{\alpha+1}>0.$$

Also,

$$(2r+2)/(2^{\alpha-1}\Pi-1)>T/2^{\alpha-1}\Pi(2^{\alpha-1}\Pi-1)+1/2^{\alpha-1}\Pi,$$

because $$2r+1>T/2^{\alpha-1}\Pi,\text{ since }r\geqq 1.$$

It remains to be shown that $T>r^2$. Now,

$$T>2^\alpha+\sigma=2^\alpha+p_1+p_2+\cdots+p_r.$$

But,

$$p_1+p_2+p_3+\cdots+p_r\geqq 3+5+7+\cdots+(2r+1)=r^2+2r.$$

Therefore,

$$n<2^{\alpha-1}\prod_1^r p_k u/(2^{\alpha-1}\prod_1^r p_k-1)+2^{\alpha-1}\epsilon/(2^{\alpha-1}-1)+1.$$

STANFORD UNIVERSITY, CALIFORNIA,
FEBRUARY 27, 1933.

CHARACTERIZATION OF SPHERICAL AND PSEUDO-SPHERICAL SETS OF POINTS.†

By Leonard M. Blumenthal ‡ and George A. Garrett.

1. *Introduction.* The n-dimensional spherical space, $S_{n,r}$, consists of the points of the surface of an $(n+1)$-dimensional sphere of radius r in a euclidean space of $n+1$ dimensions, with the distance between two points defined as the length of the shorter arc of the circle formed by the intersection of the two-dimensional plane through the two points and the center of the sphere, with the surface of the sphere. A set of points is called r-spheric (S_n) provided the set is congruent with a subset of $S_{n,r}$; while a set of $n+3$ points which is not r-spheric though each $n+2$ of the points is congruent with $n+2$ points of the $S_{n,r}$ is said to form a pseudo r-spheric $(n+3)$-tuple.

The spherical space $S_{n,r}$ is a semi-metric space, and the purpose of this paper is to obtain theorems that afford a characterization of this space among general semi-metric spaces in terms of relations between the distances of its points. In addition to these theorems, certain properties of pseudo r-spheric sets are obtained. The paper is conveniently divided into three sections.

Section I. The circle $S_{1,r}$. We denote the metric diameter of the circle by $d=\pi r$ and call a set of points d-cyclic if the set is congruent with a subset of this circle. Pseudo d-cyclic sets are sets that are not congruent with a subset of the circle, while each triple of points contained in the set is d-cyclic. Both d-cyclic and pseudo d-cyclic sets are characterized by means of distance relations expressed in determinantal form.§ The principal theorem characterizing pseudo d-cyclic sets proves that such sets are *equilateral* provided they contain more than four points, and no four of the points form a convex tripod.¶ Finally, it is shown how the three types of pseudo d-cyclic quadruples may be constructed by means of reflections in a circle.

† Presented by title at the Christmas meeting of the American Mathematical Society, December, 1932.

‡ National Research Fellow.

§ For a characterization of these sets expressed in terms of the "between-ness relation" see two papers by L. M. Blumenthal, *American Journal of Mathematics*, Vol. 54 (1932), pp. 387-396; pp. 729-738.

¶ Four points form a convex tripod if one of the points lies between each of the three pairs of points contained in the remaining three points. The point q is said to lie between the points p and r if $pq+qr=pr$; $p\neq q$, $r\neq q$.

Section II. The spherical space $S_{2,r}$. This section contains the characterization of the sphere in R_3, as well as a theorem characterizing pseudo r-spheric quintuples.

Section III. The n-dimensional spherical space $S_{n,r}$. This section obtains the necessary and sufficient conditions that $n+1$ points, $n+2$ points, $n+3$ points of a semi-metric space be congruent with a subset of the $S_{n,r}$. Since the $S_{n,r}$ is known to have the congruence order $n+3$, a semi-metric space is congruent with a subset of the $S_{n,r}$ provided each $n+3$ points of the space is congruent with $n+3$ points of the $S_{n,r}$.† Thus, the characterization of r-spheric sets is complete. In addition, it is shown that the determinant of a pseudo $(n+3)$-tuple is negative. These theorems are obtained by an induction from the cases treated in Sections I and II.

Section I. The circle $S_{1,r}$.

1. Let p_i and p_j be any two points of a semi-metric space, and denote by $\alpha_{i,j}$ the angle p_ip_j/r radians, where r is the euclidean radius of the circle of metric diameter $d=\pi r$. If $p_1, p_2, \cdots, p_n$ are n points of a semi-metric space, we denote the axisymmetric determinant

$$\begin{vmatrix} 1 & \cos\alpha_{12} & \cdots & \cos\alpha_{1n} \\ \cos\alpha_{21} & 1 & \cdots & \cos\alpha_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cos\alpha_{n1} & \cos\alpha_{n2} & \cdots & 1 \end{vmatrix}$$

by $\Delta(p_1, p_2, \cdots, p_n)$.

THEOREM 1. *Three points p_1, p_2, p_3 of a semi-metric space are d-cyclic if and only if $0 < \alpha_{i,j} \leq \pi$, $(i, j = 1, 2, 3)$, $i \neq j$, and $\Delta(p_1, p_2, p_3) = 0$.*

Evaluating the determinant of the three points, we find

$$\Delta(p_1, p_2, p_3) = 4 \sin\frac{\alpha_{12}+\alpha_{23}+\alpha_{13}}{2} \sin\frac{\alpha_{12}+\alpha_{23}-\alpha_{13}}{2}$$
$$\times \sin\frac{\alpha_{12}-\alpha_{23}+\alpha_{13}}{2} \sin\frac{-\alpha_{12}+\alpha_{23}+\alpha_{13}}{2}.$$

Since each angle is positive and at most equal to π, this expression vanishes if and only if one angle is the sum of the other two, or the sum of the three angles is 2π. Then the points p_1, p_2, p_3 are either linear or the sum of the three distances they determine equals $2d$, while each distance is at most equal

† Karl Menger, "New foundations of euclidean geometry," *American Journal of Mathematics*, Vol. 53 (1931), p. 725.

to d. But it has been shown that these are the necessary and sufficient conditions that three points be d-cyclic.† Hence the theorem follows.

Three points of a semi-metric space are said to be *circular* provided the points are congruent with three points of some circle. We state the following theorem, the proof of which is obvious:

THEOREM 2. *Three points of a semi-metric space are circular if and only if their distances satisfy the triangle inequality.*

Thus, a metric space might be defined as a semi-metric space that has each triple of its points circular.

THEOREM 3. *Four points p_1, p_2, p_3, p_4 of a semi-metric space are d-cyclic if and only if each triple is d-cyclic and* $\Delta(p_1, p_2, p_3, p_4) = 0$.

The necessity of the conditions is immediate.‡

To prove the sufficiency of the conditions, we suppose that p_1, p_2, p_3, p_4 are such that each three of the points is d-cyclic and the determinant Δ is equal to zero. We show the existence of four points, p'_1, p'_2, p'_3, p'_4, on a circle of metric diameter d which are congruent with the four given points.

At least one of the angles $\alpha_{i,j}$ is different from π, for otherwise the determinant has the value -16, contrary to the hypothesis that it vanishes. We assume the labeling so that $\alpha_{12} \neq \pi$; that is $p_1p_2 \neq d$. By hypothesis, there exist three points, say p'_1, p'_2, p'_3, and three points, say $\bar{p}_1, \bar{p}_2, \bar{p}_4$ of the circle of metric diameter d such that $p_1, p_2, p_3 \approx p'_1, p'_2, p'_3$ and $p_1, p_2, p_4 \approx \bar{p}_1, \bar{p}_2, \bar{p}_4$.§ Then $p'_1p'_2 = p_1p_2 = \bar{p}_1\bar{p}_2$, and we may make a congruent transformation of the circle into itself transforming $\bar{p}_1$ and $\bar{p}_2$ into p'_1 and p'_2 respectively. This transformation sends the point $\bar{p}_4$ into a point p'_4 which has its distances from two non-diametral points fixed and hence is uniquely determined. We now have $p_1, p_2, p_3 \approx p'_1, p'_2, p'_3$ and $p_1, p_2, p_4 \approx p'_1, p'_2, p'_4$. In order to prove the theorem, we have merely to show that $p_3p_4 = p'_3p'_4$.

---

† L. M. Blumenthal, "A complete characterization of proper pseudo d-cyclic sets of points," *American Journal of Mathematics*, Vol. 54 (1932), p. 388.

‡ Take the origin of a two-dimensional cartesian coördinate system at the center of the circle, and let A_i, B_i denote the direction cosines of the line joining the origin with the point p_i. The determinant of the four points is then easily factorable into two determinants, each of which is equal to zero. See also Lemma 1, Section II.

§ The sign $\approx$ is the symbol of congruence.

$$\Delta(x) = \begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{13} & \cos\alpha_{14} \\ \cos\alpha_{21} & 1 & \cos\alpha_{23} & \cos\alpha_{24} \\ \cos\alpha_{31} & \cos\alpha_{32} & 1 & \cos x \\ \cos\alpha_{41} & \cos\alpha_{42} & \cos x & 1 \end{vmatrix}.$$

$\Delta(x)$ does not vanish identically since the coefficient of $\cos^2 x$, namely $-\sin^2\alpha_{12}$, does not vanish. By hypothesis, a root of $\Delta(x) = 0$ is p_3p_4/r, and by the necessity of the conditions another root of the equation is $p'_3p'_4/r$. Now $\Delta(x) = 0$ has only two roots in the interval $0 < x \leq \pi$. We show that p_3p_4/r and $p'_3p'_4/r$ are double roots, and hence are equal.

We may expand $\Delta(x)$ in the form †

$$\Delta(x) = \frac{\Delta(p_1, p_2, p_3) \cdot \Delta(p_1, p_2, p_4) - \begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{13} \\ \cos\alpha_{21} & 1 & \cos\alpha_{23} \\ \cos\alpha_{41} & \cos\alpha_{42} & \cos x \end{vmatrix}^2}{1 - \cos^2\alpha_{12}}$$

and, since by Theorem 1, $\Delta(p_1, p_2, p_4)$ and $\Delta(p_1, p_2, p_3)$ have the value zero, we may write the equation in the form

$$\begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{13} \\ \cos\alpha_{21} & 1 & \cos\alpha_{23} \\ \cos\alpha_{41} & \cos\alpha_{42} & \cos x \end{vmatrix}^2 = 0.$$

Hence p_3p_4/r and $p'_3p'_4/r$ are double roots, and $p_3p_4 = p'_3p'_4$ as was to be proved.

Since the circle has the congruence order four, a set of points of a semimetric space is d-cyclic if and only if each quadruple contained in the points is d-cyclic.

We have defined a pseudo d-cyclic set of points as a set which is not d-cyclic though each triple contained in the four points is d-cyclic. The remainder of this section is devoted to a characterization of these sets. It has been shown that there exist three types of pseudo d-cyclic quadruples: the pseudo d-cyclic convex tripod, the pseudo-linear pseudo d-cyclic quadruple, and the proper pseudo d-cyclic quadruple, containing exactly three, four, and no linear triples respectively.‡

It is easily seen that if four points form a convex tripod with each triple d-cyclic, then the sum of opposite distances equals d; i. e., we may assume

† E. B. Stouffer, "Expression for a determinant in terms of five minors," *American Mathematical Monthly*, Vol. 39 (1932), p. 165.

‡ L. M. Blumenthal, *loc. cit.*

the labeling of the points so that $\cos\alpha_{24} = -\cos\alpha_{13}$, $\cos\alpha_{34} = -\cos\alpha_{12}$, $\cos\alpha_{14} = -\cos\alpha_{23}$, and $\cos\alpha_{23} = \cos(\alpha_{12} + \alpha_{13})$. The determinant of the four points then takes the form

$$\text{(I)} \qquad \Delta(p_1, p_2, p_3, p_4) = \begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{13} & \cos\alpha_{23} \\ \cos\alpha_{12} & 1 & \cos\alpha_{23} & \cos\alpha_{13} \\ \cos\alpha_{13} & \cos\alpha_{23} & 1 & \cos\alpha_{12} \\ \cos\alpha_{23} & \cos\alpha_{13} & \cos\alpha_{12} & 1 \end{vmatrix}.$$

If p_1, p_2, p_3, p_4 form a pseudo-linear pseudo d-cyclic quadruple, then opposite distances are equal and no two points are diametral. We may assume the labeling so that $\cos\alpha_{24} = \cos\alpha_{13}$, $\cos\alpha_{34} = \cos\alpha_{12}$, $\cos\alpha_{14} = \cos\alpha_{23}$, and $\cos\alpha_{23} = \cos(\alpha_{12} + \alpha_{13})$. From these relations we see that the determinant of the four points is identical with (I) above. Similarly, it is seen that the determinant of four points forming a proper pseudo d-cyclic quadruple is given by (I) with the cosines of opposite angles equal and the labeling assumed so that $\cos\alpha_{23} = \cos(\alpha_{12} + \alpha_{13})$.

Summarizing the above remarks, we obtain the following lemma:

LEMMA. *If four points p_1, p_2, p_3, p_4 of a semi-metric space form a pseudo d-cyclic quadruple, their determinant is*

$$\begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{13} & \cos\alpha_{23} \\ \cos\alpha_{12} & 1 & \cos\alpha_{23} & \cos\alpha_{13} \\ \cos\alpha_{13} & \cos\alpha_{23} & 1 & \cos\alpha_{12} \\ \cos\alpha_{23} & \cos\alpha_{13} & \cos\alpha_{12} & 1 \end{vmatrix}$$

where $\cos\alpha_{23} = \cos(\alpha_{12} + \alpha_{13})$ and none of the angles appearing in the determinant has the value π.

Developing this determinant, we find that

$$\Delta(p_1, p_2, p_3, p_4) = -4\sin^2\alpha_{12}\sin^2\alpha_{23}\sin^2\alpha_{13}.$$

We have, then, the following theorems:

THEOREM 4. *The determinant of a pseudo d-cyclic quadruple is negative.*

THEOREM 5. *If all four third-order principal minors of the determinant $\Delta(p_1, p_2, p_3, p_4)$ vanish, and $0 < \alpha_{i,j} \leq \pi$, $(i, j = 1, 2, 3, 4)$, $i \neq j$, then the determinant either vanishes or has the value $-4\sin^2\alpha_{12}\sin^2\alpha_{23}\sin^2\alpha_{13}$. In the latter case, each angle is less than π, the squares of the cosines of opposite angles are equal, and the labelling may be assumed so that $\cos\alpha_{23} = \cos(\alpha_{12} + \alpha_{13})$.*

Three corollaries may be stated giving the necessary and sufficient condition that four points form a pseudo d-cyclic quadruple of one of the three types.

Construction of pseudo d-cyclic quadruples. Let p'_1, p'_2, p'_3 be three non-linear points which are d-cyclic. Reflect p'_3 in the diameters through p'_1 and p'_2, obtaining p^*_3 and p^{**}_3 respectively. Let p'_4 and $\bar{p}'_4$ be the two points of the circle equidistant from p^*_3 and p^{**}_3, with p'_4 the mid-point of the shorter arc joining the last two points. Consider four points p_1, p_2, p_3, p_4 with distances defined as follows: $p_1p_2 = p'_1p'_2$, $p_1p_3 = p'_1p'_3$, $p_2p_3 = p'_2p'_3$ $p_1p_4 = p'_1p'_4$, $p_2p_4 = p'_2p'_4$, $p_3p_4 = p^*_3p'_4 = p^{**}_3p'_4$. It is readily seen that the four points form a convex tripod with each triple d-cyclic, and the point p_4 between each of the three pairs of points contained in p_1, p_2, p_3.

Consider now four points p_1, p_2, p_3, p_4 in which the distances p_1p_2, p_1p_3, p_2p_3 are defined as above, while

$$p_1p_4 = p'_1\bar{p}'_4, \quad p_2p_4 = p'_2\bar{p}'_4, \quad p_3p_4 = p^*_3\bar{p}'_4 = p^{**}_3\bar{p}'_4.$$

Then each triple is d-cyclic, not linear, and opposite distances are equal. The points form a proper pseudo d-cyclic quadruple. If p'_1, p'_2, p'_3 are chosen equilateral, a proper pseudo d-cyclic quadruple which is equilateral may be constructed in this way.

Let, now, the points p'_1, p'_2, p'_3 be chosen linear, with no two diametral, and determine the points p'_4 and $\bar{p}'_4$ as above. Then each triple is d-cyclic and linear, opposite distances are equal, and the four points form a pseudo d-cyclic pseudo-linear quadruple.

Pseudo d-cyclic sets no four points of which form a convex tripod are called *regular.* In order to establish the theorem characterizing such sets which contain more than four points we prove two lemmas concerning the fifth order determinant $\Delta(p_1, p_2, p_3, p_4, p_5)$. We make the following hypotheses:

(a). Each angle of the determinant is positive and at most equal to π.

(b). Each third-order principal minor vanishes.

(c). At least one fourth-order principal minor does not vanish.

(d). If $\Delta(p_i, p_j, p_k, p_l)$ is any non-vanishing fourth-order principal minor, it is possible so to label the elements of the minor that $\cos \alpha_{ij} + \cos \alpha_{kl} \neq 0$.†

LEMMA 1. *$\Delta(p_1, p_2, p_3, p_4, p_5)$ does not contain exactly one non-vanishing fourth-order principal minor.*

† This hypothesis excludes the possibility of any four points forming a pseudo d-cyclic convex tripod.

From hypothesis (c) we may assume that $\Delta(p_1, p_2, p_3, p_4)$ does not vanish. Then be Theorem 5, $\Delta(p_1, p_2, p_3, p_4)$ is negative and none of the six angles contained in this minor has the value π. From hypothesis (d) the four points p_1, p_2, p_3, p_4 do not form a convex tripod and hence the cosines of opposite angles are equal. We suppose that the four remaining fourth-order principal minors all vanish, and we show that this assumption leads to a contradiction.

Expanding each of these four minors, and writing $a = \cos\alpha_{12} = \cos\alpha_{34}$, $b = \cos\alpha_{13} = \cos\alpha_{24}$, $c = \cos\alpha_{23} = \cos\alpha_{14}$ we obtain

$$\begin{aligned}
&(ac-b)\cos\alpha_{15} + (ab-c)\cos\alpha_{25} + (1-a^2)\cos\alpha_{35} &&= 0\\
&(ab-c)\cos\alpha_{15} + (ac-b)\cos\alpha_{25} + (1-a^2)\cos\alpha_{45} &&= 0\\
&(ab-c)\cos\alpha_{15} + (bc-a)\cos\alpha_{35} + (1-b^2)\cos\alpha_{45} &&= 0\\
&(ac-b)\cos\alpha_{25} + (bc-a)\cos\alpha_{35} + (1-c^2)\cos\alpha_{45} &&= 0.
\end{aligned}$$

Now $\cos\alpha_{15}$, $\cos\alpha_{25}$, $\cos\alpha_{35}$, $\cos\alpha_{45}$ are not all zero. In fact we get a contradiction by supposing that two of them, say $\cos\alpha_{15}$ and $\cos\alpha_{25}$ are zero. For if so, then $\alpha_{15} = \alpha_{25} = \pi/2$, and since $\Delta(p_1, p_2, p_5) = 0$, then α_{12} must have the value 0 or π, which is impossible. The four equations cannot be satisfied, then, unless

$$\begin{vmatrix}
ac-b, & ab-c, & 1-a^2, & 0\\
ab-c, & ac-b, & 0\ \ , & 1-a^2\\
ab-c, & 0\ \ , & bc-a, & 1-b^2\\
0\ \ , & ac-b, & bc-a, & 1-c^2
\end{vmatrix} = 0.$$

Now since $\Delta(p_1, p_2, p_3) = 0$, by Theorem 1 either the sum of the three angles α_{12}, α_{23}, α_{13} equals 2π or one angle is the sum of the other two. Thus, two cases present themselves.

Case 1. $\alpha_{12} + \alpha_{23} + \alpha_{13} = 2\pi$.

In this case the above equation is readily put in the form

$$\begin{vmatrix}
0 & \sin\alpha_{12} & \sin\alpha_{13} & \sin\alpha_{23}\\
\sin\alpha_{12} & 0 & \sin\alpha_{23} & \sin\alpha_{13}\\
\sin\alpha_{13} & \sin\alpha_{23} & 0 & \sin\alpha_{12}\\
\sin\alpha_{23} & \sin\alpha_{13} & \sin\alpha_{12} & 0
\end{vmatrix} = 0$$

and evaluation of the determinant yields

$$\begin{aligned}
&(\sin\alpha_{12} + \sin\alpha_{23} + \sin\alpha_{13})(\sin\alpha_{12} + \sin\alpha_{23} - \sin\alpha_{13})\\
&\quad\times(\sin\alpha_{12} - \sin\alpha_{23} + \sin\alpha_{13})(-\sin\alpha_{12} + \sin\alpha_{23} + \sin\alpha_{13}) = 0.
\end{aligned}$$

But it is readily shown that no one of the factors of the above expression can vanish. Hence, in this case we have obtained the desired contradiction.

Case 2. We may assume that $\alpha_{13} = \alpha_{12} + \alpha_{23}$. In this case the determinant obtained is merely the negative of the one obtained in Case 1, and it is again easily shown not to vanish.

Thus the assumption that exactly four of the fourth-order principal minors of the determinant $\Delta(p_1, p_2, p_3, p_4, p_5)$ vanish is seen to lead to a contradiction, and the theorem is proved.

LEMMA 2. *No fourth-order principal minor of the determinant* $\Delta(p_1, p_2, p_3, p_4, p_5)$ *vanishes.*

By the preceding lemma, there are at least two non-vanishing fourth-order principal minors. We assume the labeling so that $\Delta(p_1, p_2, p_3, p_4)$ and $\Delta(p_1, p_2, p_3, p_5)$ do not vanish. Then none of the nine angles contained in these two minors has the value π, while we have

$$\begin{aligned} &\cos\alpha_{12} = \cos\alpha_{34} = \cos\alpha_{35} \\ (1)\qquad &\cos\alpha_{13} = \cos\alpha_{24} = \cos\alpha_{25} \\ &\cos\alpha_{23} = \cos\alpha_{14} = \cos\alpha_{15}. \end{aligned}$$

Suppose that any other fourth-order principal minor, say $\Delta(p_1, p_2, p_4, p_5)$ vanishes. Then from the expansion of this minor we have

$$\begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{14} \\ \cos\alpha_{21} & 1 & \cos\alpha_{24} \\ \cos\alpha_{15} & \cos\alpha_{25} & \cos\alpha_{45} \end{vmatrix} = 0.$$

Applying (1), we may write this equation in the form

$$\begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{14} \\ \cos\alpha_{21} & 1 & \cos\alpha_{24} \\ \cos\alpha_{14} & \cos\alpha_{24} & \cos\alpha_{45} \end{vmatrix} = 0.$$

Consider the function $\phi(x)$ defined as follows:

$$\phi(x) \equiv \begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{14} \\ \cos\alpha_{21} & 1 & \cos\alpha_{24} \\ \cos\alpha_{14} & \cos\alpha_{24} & x \end{vmatrix}.$$

Since $\Delta(p_1, p_2, p_4) = 0$, one root of $\phi(x) = 0$ is $x = 1$. Since the coefficient of x in the equation does not vanish, $x = 1$ is the only root of the equation. Then we must have $\cos\alpha_{45} = 1$; i. e., $\alpha_{45} = 0$, which is impossible. Hence the lemma is proved.

THEOREM 6. *Each element of the determinant* $\Delta(p_1, p_2, p_3, p_4, p_5)$ *out-*

side of the principal diagonal has the value $-1/2$, *and the value of the determinant is* $-(3/2)^4$.

Since no fourth-order principal minor of the determinant vanishes, and by hypothesis (d) $\cos\alpha_{ij}+\cos\alpha_{kl}\neq 0$, where α_{ij} and α_{kl} are opposite angles in a non-vanishing fourth-order principal minor, then opposite angles occurring in each fourth-order principal minor are equal. We obtain, then, that each of the ten angles are equal, and since each third-order principal minor is zero, each angle equals $2\pi/3$.

Theorem 6 is the determinant form of the theorem:

THEOREM 7. *A regular pseudo d-cyclic quintuple is equilateral.*

Applying mathematical induction we obtain the more general theorem: †

THEOREM 8. *A regular pseudo d-cyclic set containing more than four points is equilateral.*

This theorem is equivalent to the following interesting theorem on determinants:

THEOREM 9. *If* $\Delta(p_1, p_2, \cdots, p_n)$, $n>4$, *is such that*

(a). *Each angle is positive and at most equal to* π.

(b). *Each third-order principal minor vanishes.*

(c). *At least one fourth-order principal minor does not vanish.*

(d). *If* $\Delta(p_i, p_j, p_k, p_l)$ *is any non-vanishing fourth-order principal minor, it is possible so to arrange the labelling of the elements of the minor that* $\cos\alpha_{ij}+\cos\alpha_{kl}\neq 0$.

Then each angle contained in the determinant has the value $2\pi/3$ *and the determinant has the value* $-1/2\,(3/2)^{n-1}(n-3)$.

Section II. The sphere $S_{2,r}$.

In this section we characterize r-spheric ‡ and pseudo r-spheric sets by means of theorems similar to those characterizing the sets treated in Section I. The proofs of the necessity of the conditions imposed upon the points are considerably shortened by means of the following lemma which we prove at once for the $(n-1)$-dimensional spherical space $S_{n-1,r}$.

We define the angles α_{ij} and the determinant $\Delta(p_1, p_2, \cdots, p_n)$ as in Section I. Let O denote the center of the n-dimensional sphere whose surface

† L. M. Blumenthal, *loc. cit.*

‡ Throughout this section the term "r-spheric" has reference only to the S_2 of radius r.

is the space $S_{n-1,r}$, and $v(p_1, p_2, \cdots, p_n, O)$ the volume of the simplex determined by the $n+1$ points $p_1, p_2, \cdots, p_n, O$, where $p_1, p_2, \cdots, p_n$ are points of $S_{n-1,r}$.

LEMMA 1. $\Delta(p_1, p_2, \cdots, p_n) = (n!/r^n)^2 v^2(p_1, p_2, \cdots, p_n, O)$.

Let (ij) denote the square of the euclidean distance of the points p_i, p_j. We have, by a well-known theorem,

$$v^2(p_1, p_2, \cdots, p_n, O) = [(-1)^{n+1}/(n!)^2 2^n] D(p_1, p_2, \cdots, p_n, O),$$

where

$$D(p_1, p_2, \cdots, p_n, O) = \begin{vmatrix} 0 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & (1n) & r^2 \\ 1 & (21) & \cdots & (2n) & r^2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 1 & (n1) & \cdots & 0 & r^2 \\ 1 & r^2 & \cdots & r^2 & 0 \end{vmatrix}.$$

Subtracting the last row of this determinant from each preceding row except the first, subtracting the last column from each preceding column, and substituting $(ij) = 2r^2(1 - \cos \alpha_{ij})$, we find

$$D(p_1, p_2, \cdots, p_n, O) = (-1)^{n+1}(2r^2)^n \Delta(p_1, p_2, \cdots, p_n).$$

Putting this value for D in the expression for v^2 written above, we obtain the lemma.

LEMMA 2. *If $p_1, p_2, \cdots, p_n$, $(n > 3)$ are congruent with n points of the $S_{2,r}$ then $\Delta(p_1, p_2, \cdots, p_n) = 0$.*

This follows immediately from the above lemma, for the determinant Δ is evidently a congruence invariant, while the simplex formed by the points is degenerate and has zero volume.

If $n = 3$, then $v^2(p_1, p_2, p_3, O)$ equals zero if the points p_1, p_2, p_3, O lie in a plane, and is positive otherwise. Since these points lie in a plane if and only if p_1, p_2, p_3 lie on a great circle of the sphere, we have the following lemma.

LEMMA 3. *If p_1, p_2, p_3 are r-spheric (S_2) and not d-cyclic, then $\Delta(p_1, p_2, p_3)$ is positive.*

THEOREM 10. *Three points p_1, p_2, p_3 of a semi-metric space are r-spheric if and only if $0 < \alpha_{ij} \leq \pi$, $(i, j = 1, 2, 3)$, $i \neq j$, and $\Delta(p_1, p_2, p_3) \geq 0$.*

The necessity of the conditions follows immediately from Lemma 3 and Theorem 1. To prove the sufficiency of the conditions, we consider two cases.

Case 1. $\Delta(p_1, p_2, p_3) = 0$. Then the three points are d-cyclic, by Theorem 1, and hence are congruent with three points of a great circle of the sphere.

Case 2. $\Delta(p_1, p_2, p_3) > 0$. It is sufficient to show that the three angles $\alpha_{12}, \alpha_{23}, \alpha_{13}$, satisfy the triangle inequality, and have a sum less than 2π.

Each of the three angles cannot have the value π, for if so, the determinant has the value -4, contrary to hypothesis. Hence the sum of the angles is less than 3π. Now the sum of the angles does not equal 2π, since in this case the determinant $\Delta(p_1, p_2, p_3)$ would be zero. We suppose, then, that $2\pi < \alpha_{12} + \alpha_{23} + \alpha_{13} < 3\pi$, and obtain a contradiction.

We have $\Delta(p_1, p_2, p_3) = 4 \sin A \cdot \sin B \cdot \sin C \cdot \sin D$, where

$$A = \tfrac{1}{2}(\alpha_{12} + \alpha_{23} + \alpha_{13}), \qquad B = \tfrac{1}{2}(\alpha_{12} + \alpha_{23} - \alpha_{13}),$$
$$C = \tfrac{1}{2}(\alpha_{12} - \alpha_{23} + \alpha_{13}), \qquad D = \tfrac{1}{2}(-\alpha_{12} + \alpha_{23} + \alpha_{13}).$$

According to the supposition made above, the angle A lies between π and $3\pi/2$. It follows, then, that each of the angles B, C, D are positive and less than π. Then their sines are positive, while the sine of A is negative. Hence $\Delta(p_1, p_2, p_3)$ is negative, which contradicts the hypothesis that the determinant of the three points is positive. Hence $0 < \alpha_{12} + \alpha_{23} + \alpha_{13} < 2\pi$.

Since $\Delta(p_1, p_2, p_3)$ is positive and $\sin A$ is positive, the product of the three factors, $\sin B$, $\sin C$, $\sin D$ must be positive. Hence, either all three of the factors are positive, or two of them are negative and one is positive. It is readily shown that this latter case cannot occur. Thus, each factor is positive; i. e., each angle is positive, and the angles α_{12}, α_{23}, α_{13}, satisfy the triangle inequality.

THEOREM 11. *Four points p_1, p_2, p_3, p_4 of a semi-metric space are r-spheric if and only if each three of the points is r-spheric and $\Delta(p_1, p_2, p_3, p_4) = 0$.*

If the four points are r-spheric then each three is r-spheric and, by Lemma 2, $\Delta(p_1, p_2, p_3, p_4)$ equals zero.

To prove the sufficiency, we suppose that each triple is r-spheric and that $\Delta(p_1, p_2, p_3, p_4) = 0$. We show the existence of four points of the sphere congruent to the given four points.

The proof is immediate if each triple is d-cyclic; for then, by Theorem 3, the four points are d-cyclic. We suppose, then, that at least one triple, p_1, p_2, p_3 is not d-cyclic. Then no one of the angles $\alpha_{12}, \alpha_{23}, \alpha_{13}$, is equal to π. By hypothesis there exist points p'_1, p'_2, p'_3 and $\bar{p}_1, \bar{p}_2, \bar{p}_4$ of the sphere such that

$$p_1, p_2, p_3 \approx p'_1, p'_2, p'_3; \qquad p_1, p_2, p_4 \approx \bar{p}_1, \bar{p}_2, \bar{p}_4.$$

11

Then $p_1p_2 = p'_1p'_2 = \bar{p}_1\bar{p}_2$, and we may make a congruent transformation of the sphere into itself transforming $\bar{p}_1$ into p'_1 and $\bar{p}_2$ into p'_2. This transformation sends $\bar{p}_4$ into some point, say p^*_4, such that $p'_1p^*_4 = p_1p_4$ and $p'_2p^*_4 = p_2p_4$. Since the point p^*_4 has its distances from the non-diametral points p'_1 and p'_2 determined, there are at most two such points p^*_4. Two cases present themselves.

Case 1. *The points p_1, p_2, p_4 are not d-cyclic.* In this case there are two possible positions on the sphere for the point p^*_4. We denote them by p_4^I and p_4^{II}. Then these two points are reflections of each other in the plane through p'_1, p'_2, and the center of the sphere. Since p_1, p_2, p_3 are not d-cyclic, the point p'_3 is not on this plane and therefore $p'_3p_4^I \neq p'_3p_4^{II}$. We now have

$$p_1, p_2, p_3 \approx p'_1, p'_2, p'_3; \qquad p_1, p_2, p_4 \approx p'_1, p'_2, p_4^I \approx p'_1, p'_2, p_4^{II}.$$

Hence, in order to prove the theorem we have only to show that $p_3p_4 = p'_3p_4^I$ or $p_3p_4 = p'_3p_4^{II}$. In order to do this we define the function

$$\Delta(x) = \begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{13} & \cos\alpha_{14} \\ \cos\alpha_{21} & 1 & \cos\alpha_{23} & \cos\alpha_{24} \\ \cos\alpha_{31} & \cos\alpha_{32} & 1 & \cos x \\ \cos\alpha_{41} & \cos\alpha_{42} & \cos x & 1 \end{vmatrix}.$$

Now $\Delta(x)$ is not identically zero, for the coefficient of $\cos^2 x$ does not vanish. By the necessity of the theorem, $\Delta(x) = 0$ has the two unequal roots $p'_3p_4^I/r$ and $p'_3p_4^{II}/r$, while by hypothesis one root of the equation is p_3p_4/r. Since the equation has only two roots in the interval $0 < x \leq \pi$, we have $p_3p_4 = p'_3p_4^I$ or $p_3p_4 = p'_3p_4^{II}$, as was to be proved.

Case 2. *The points p_1, p_2, p_4 are d-cyclic.* In this case the point p^*_4 is uniquely determined. We denote the point by p'_4, and shall show that $p_1, p_2, p_3, p_4 \approx p'_1, p'_2, p'_3, p'_4$.

We have $p_1, p_2, p_3 \approx p'_1, p'_2, p'_3$ and $p_1, p_2, p_4 \approx p'_1, p'_2, p'_4$, and we need only show that $p_3p_4 = p'_3p'_4$. To do this, we observe that since p_1, p_2, p_4 are d-cyclic, $\Delta(p_1, p_2, p_4) = 0$, and hence the equation $\Delta(x) = 0$ has a double root as its only roots in the interval $0 < x \leq \pi$.† Hence, $p_3p_4 = p'_3p'_4$ and or $p_3p_4 = p'_3p_4^{II}$, as was to be proved.

THEOREM 12. *Five points p_1, p_2, p_3, p_4, p_5, of a semi-metric are r-spheric if and only if each four of the points is r-spheric and $\Delta(p_1, p_2, p_3, p_4, p_5) = 0$.*

The necessity of the conditions is evident. Now by Theorem 11 each

† This is shown in the same manner as in Theorem 3, Section I.

fourth-order principal minor of Δ vanishes. If each triple of the points is d-cyclic, then by Theorem 3 each four of the points is d-cyclic, and since the circle has the congruence order four, the five points are d-cyclic.

We suppose that at least one triple, p_1, p_2, p_3 is not d-cyclic. Then $\Delta(p_1, p_2, p_3)$ is positive. By hypothesis there exist points of the sphere such that

$$p_1, p_2, p_3, p_4 \approx p'_1, p'_2, p'_3, p'_4 \quad \text{and} \quad p_1, p_2, p_3, p_5 \approx \bar{p}_1, \bar{p}_2, \bar{p}_3, \bar{p}_5.$$

We may make a congruent transformation of the sphere into itself transforming $\bar{p}_1$ into p'_1, $\bar{p}_2$ into p'_2, and $\bar{p}_3$ into p'_3. Since p_1, p_2, p_3 are not d-cyclic, the transform of the point $\bar{p}_5$ is uniquely determined. Denote it by p'_5. We wish to show that $p_4p_5 = p'_4p'_5$.

Again we consider a function $\Delta(x)$ obtained from $\Delta(p_1, p_2, p_3, p_4, p_5)$ by replacing α_{45} by x. This function is quadratic in $\cos x$, since the coefficient of $\cos^2 x$, namely $-\Delta(p_1, p_2, p_3)$, does not vanish. By the necessity of the conditions of the theorem the equation $\Delta(x) = 0$ has a root $x = p'_4p'_5/r$, while by hypothesis $x = p_4p_5/r$ is also a root.

Since, by Theorem 11, $\Delta(p_1, p_2, p_3, p_4) = 0$, we have upon expanding $\Delta(x)$ by the method of Stouffer,

$$\Delta(x) = \frac{-\begin{vmatrix} 1 & \cos\alpha_{12} & \cos\alpha_{13} & \cos\alpha_{14} \\ \cos\alpha_{21} & 1 & \cos\alpha_{23} & \cos\alpha_{24} \\ \cos\alpha_{31} & \cos\alpha_{32} & 1 & \cos\alpha_{34} \\ \cos\alpha_{51} & \cos\alpha_{52} & \cos\alpha_{53} & \cos x \end{vmatrix}^2}{\Delta(p_1, p_2, p_3)}.$$

Hence the roots of $\Delta(x) = 0$ exhibited above are double roots and therefore $p_4p_5 = p'_4p'_5$, as was to be proved.

Theorems 10, 11, and 12 characterize r-spheric triples, quadruples and quintuples. Since the sphere has the congruence order five, the characterization of r-spheric sets for the S_2 is complete. The remainder of this section is devoted to pseudo r-spheric quintuples.

THEOREM 13. *If five points form a pseudo r-spheric quintuple, their determinant is negative.*

From the definition of a pseudo r-spheric quintuple, and from Theorem 11, each fourth-order principal minor of $\Delta(p_1, p_2, p_3, p_4, p_5)$ vanishes. Then at least one third-order principal minor does not vanish; that is, at least one triple is not d-cyclic, for otherwise each four of the points would be d-cyclic and consequently the five points would not be pseudo r-spheric.

We assume the labelling so that p_1, p_2, p_3 are not d-cyclic; that is $\Delta(p_1, p_2, p_3)$ is positive. Evaluating the determinant of the five points we obtain $\Delta(\alpha_{45})$ by substituting α_{45} for x in the expression for $\Delta(x)$ given in Theorem 12. Therefore the determinant is negative or zero. But if the determinant is zero, the points are r-spheric, by the preceding theorem. Hence $\Delta(p_1, p_2, p_3, p_4, p_5)$ is negative.

THEOREM 14. *None of the triples contained in a pseudo r-spheric quintuple is d-cyclic while two points are spherical isogonal conjugates with respect to the other three.*

The proof of this theorem as well as additional theorems concerning the structure of pseudo r-spheric quintuples will appear in a forthcoming paper.

Section III. The n-dimensional space $S_{n,r}$.

The $S_{n,r}$ is characterized among general semi-metric spaces by means of the following four theorems:

THEOREM I_n. *A necessary and sufficient condition that $n+1$ points of a semi-metric space be congruent with $n+1$ points of the $S_{n,r}$ is that each n of the points be congruent with n points of the $S_{n-1,r}$, and that the determinant of the $n+1$ points be positive or zero.*

THEOREM II_n. *$n+2$ points of a semi-metric space are congruent with $n+2$ points of the $S_{n,r}$ if and only if each $n+1$ of the points is congruent with $n+1$ points of the $S_{n,r}$ and the determinant of the points is equal to zero.*

THEOREM III_n. *$n+3$ points of a semi-metric space are congruent with $n+3$ points of the $S_{n,r}$ if and only if each $n+2$ of the points is congruent with $n+2$ points of the $S_{n,r}$ and the determinant of the $n+3$ points is zero.*

THEOREM IV_n. *If $n+3$ points form a pseudo $(n+3)$-tuple, their determinant is negative.*

Since the S_n has the congruence order $n+3$, the characterization is complete.

By Lemma 1, Section II, we have

$$\Delta(p_1, p_2, \cdots, p_n) = (n!/r^n)^2\, v^2(p_1, p_2, \cdots, p_n, O);$$

where the points are congruent with n points of the $S_{n-1,r}$. Whence, the determinant of m points is zero if the points are congruent with m points of the $S_{n,r}$ and m exceeds $n+1$. Also, it is evident that the determinant of

$n+1$ points which are congruent with $n+1$ points of the $S_{n,r}$ but are not congruent with $n+1$ points of the $S_{n-1,r}$ is positive.

A lemma that we shall use several times in this section is the following:

LEMMA 1. *If $p_1, p_2, \cdots, p_{n+1}$ are congruent to $n+1$ points of the $S_{n,r}$, but are not congruent to $n+1$ points of the $S_{n-1,r}$, then at least one n-tuple contained in these $n+1$ points is congruent to n points of the $S_{n-1,r}$ and is not congruent to n points of the $S_{n-2,r}$.*

Suppose the contrary. Then each n-tuple contained in the $n+1$ points is congruent to n points of the $S_{n-2,r}$. Then, the determinant of the $n+1$ points has each of its principal minors of order n vanishing. But the determinant itself is positive. Hence at least one principal minor of order $(n-1)$ does not vanish, for if each of these minors were to vanish also, the determinant would be equal to zero.

We may assume the labelling so that the minor $\Delta(p_1, p_2, \cdots, p_{n-1})$ does not vanish. The points $p_1, p_2, \cdots, p_{n-1}$ are congruent with $n-1$ points of the $S_{n-2,r}$ for we have assumed that each n points contained in the given $n+1$ points are congruent with n points of the $S_{n-2,r}$. Therefore the determinant $\Delta(p_1, p_2, \cdots, p_{n-1})$ is positive.

Evaluating the determinant $\Delta(p_1, p_2, \cdots, p_n, p_{n+1})$ by the usual method we have

$$\Delta(p_1, p_2, \cdots, p_n, p_{n+1}) = \frac{-\begin{vmatrix} 1 & \cos\alpha_{12} & \cdots & \cos\alpha_{1,n-1} & \cos\alpha_{1,n} \\ \cos\alpha_{21} & 1 & \cdots & \cos\alpha_{2,n-1} & \cos\alpha_{2,n} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cos\alpha_{n-1,1} & \cos\alpha_{n-1,2} & \cdots & 1 & \cos\alpha_{n-1,n} \\ \cos\alpha_{n+1,1} & \cos\alpha_{n+1,2} & \cdots & \cos\alpha_{n+1,n-1} & \cos\alpha_{n+1,n} \end{vmatrix}^2}{\Delta(p_1, p_2, \cdots, p_{n-1})}.$$

Hence $\Delta(p_1, p_2, \cdots, p_n, p_{n+1})$ is negative or zero, which contradicts the hypothesis that the points $p_1, p_2, \cdots, p_n, p_{n+1}$ are congruent with $n+1$ points of the $S_{n,r}$ and not congruent with $n+1$ points of the $S_{n-1,r}$. Hence the lemma is proved.

The necessity of the conditions stated in Theorems I_n, II_n and III_n follows immediately from the formula connecting the determinant Δ with the volume of a simplex.

We establish the sufficiency of the conditions by means of mathematical induction. Theorem I_1 is obviously true, while Theorems II_1 and III_1, as well as Theorem IV_1 are proved in Section I. The same theorems for $n=2$ are proved in Section 2. We shall assume the truth of all four theorems for $n=k-2$ and $n=k-1$, and show that all four theorems are true when

$n = k$. In order to prove the four theorems, we have only to prove Theorems I'_k, II'_k, III'_k, and IV'_k where the "primes" are used to denote the parts of the theorems which state that the conditions are sufficient.

THEOREM I'_k. *If $k+1$ points $p_1, p_2, \cdots, p_{k+1}$ of a semi-metric space are such that each k points is congruent to k points of the $S_{k-1,r}$ and $\Delta(p_1, p_2, \cdots, p_{k+1}) \geqq 0$, then the $k+1$ points are congruent to $k+1$ points of the $S_{k,r}$.*

If $\Delta(p_1, p_2, \cdots, p_{k+1}) = 0$, then the $k+1$ points are congruent to $k+1$ points of the $S_{k-1,r}$ by Theorem II_{k-1}. Hence the $k+1$ points are congruent to $k+1$ points of the $S_{k,r}$. We now have only to prove the theorem true when $\Delta(p_1, p_2, \cdots, p_{k+1})$ is positive. We shall treat two separate cases.

Case I. Each k points is congruent to k points of the $S_{k-2,r}$. In this case the $k+1$ points either are congruent to $k+1$ points of the $S_{k-2,r}$ or the $k+1$ points form a pseudo $(k+1)$-tuple. If the $k+1$ points are congruent to $k+1$ points of the $S_{k-2,r}$, then $\Delta(p_1, p_2, \cdots, p_{k+1})$ equals zero, by Theorem III_{k-2}. If the $k+1$ points form a pseudo $(k+1)$-tuple, then $\Delta(p_1, p_2, \cdots, p_{k+1})$ is negative, by Theorem IV_{k-2}. Both cases are impossible, for we assume that $\Delta(p_1, p_2, \cdots, p_{k+1})$ is positive.

Case II. At least one k-tuple is not congruent to k points of the $S_{k-2,r}$. We shall assume the labelling so that $p_1, p_2, \cdots, p_k$ are not congruent to k points of the $S_{k-2,r}$. Since, however, the k points $p_1, p_2, \cdots, p_k$ are congruent to k points of the $S_{k-1,r}$, then by Lemma 1 at least one $(k-1)$-tuple contained in these k points is congruent to $k-1$ points of the $S_{k-2,r}$ and is not congruent to $k-1$ points of the $S_{k-3,r}$. We may assume the labelling so that $p_1, p_2, \cdots, p_{k-1}$ are congruent to $k-1$ points of the $S_{k-2,r}$ and are not congruent to $k-1$ points of the $S_{k-3,r}$. Then,

$$(1) \qquad \Delta(p_1, p_2, \cdots, p_{k-1}) > 0.$$

Now we may expand $\Delta(p_1, p_2, \cdots, p_{k+1})$ in the form

$$\Delta(p_1, p_2, \cdots, p_{k+1}) = \frac{\Delta(p_1, p_2, \cdots, p_k)\Delta(p_1, p_2, \cdots, p_{k-1}, p_{k+1})}{\Delta(p_1, p_2, \cdots, p_{k-1})}$$

$$-\frac{\begin{vmatrix} 1 & \cos\alpha_{12} & \cdots \cos\alpha_{1,k-1} & \cos\alpha_{1,k} \\ \cos\alpha_{21} & 1 & \cdots \cos\alpha_{2,k-1} & \cos\alpha_{2,k} \\ \cdot & \cdot & \cdots & \cdot \quad\quad \cdot \\ \cos\alpha_{k-1,1} & \cos\alpha_{k-1,2} & \cdots \quad 1 & \cos\alpha_{k-1,k} \\ \cos\alpha_{k+1,1} & \cos\alpha_{k+1,2} & \cdots \cos\alpha_{k+1,k-1} & \cos\alpha_{k+1,k} \end{vmatrix}^2}{\Delta(p_1, p_2, \cdots, p_{k-1})}$$

from which we may conclude that $p_1, p_2, \cdots, p_{k-1}, p_{k+1}$ are not congruent to k points of the $S_{k-2,r}$; for otherwise $\Delta(p_1, p_2, \cdots, p_{k-1}, p_{k+1})$ equals zero and $\Delta(p_1, p_2, \cdots, p_{k+1})$ is not positive.

By hypothesis there exist k points $p'_1, p'_2, \cdots, p'_k$ of the $S_{k-1,r}$ and k points $\bar{p}_1, \bar{p}_2, \cdots, \bar{p}_{k-1}, \bar{p}_{k+1}$ of the $S_{k-1,r}$ such that

$$p_1, p_2, \cdots, p_{k-1}, p_k \approx p'_1, p'_2, \cdots, p'_{k-1}, p'_k \quad \text{and}$$
$$p_1, p_2, \cdots, p_{k-1}, p_{k+1} \approx \bar{p}_1, \bar{p}_2, \cdots, \bar{p}_{k-1}, \bar{p}_{k+1}.$$

We may make a congruent transformation of the $S_{k-1,r}$ into itself such that $\bar{p}_1$ goes into p'_1, $\bar{p}_2$ goes into $p'_2, \cdots \bar{p}_{k-1}$ goes into p'_{k-1}. This transformation sends $\bar{p}_{k+1}$ into some point p^*_{k+1}. We now have

$$p_1, p_2, \cdots, p_{k-1}, p_{k+1} \approx p'_1, p'_2, \cdots, p'_{k-1}, p^*_{k+1}.$$

The point p^*_{k+1} has its distances from the $k-1$ points $p'_1, p'_2, \cdots, p'_{k-1}$ determined. Since $p_1, p_2, \cdots, p_{k-1}, p_{k+1}$ are not congruent to k points of the $S_{k-2,r}$, the points $p'_1, p'_2, \cdots, p'_{k-1}, p^*_{k+1}$ are not on an $S_{k-2,r}$. Hence, there are exactly two distinct possible positions for the point p^*_{k+1} on the $S_{k-1,r}$ containing $p'_1, p'_2, \cdots, p'_{k-1}$. We shall denote these two points by p^I_{k+1} and p^{II}_{k+1}. These points are images of each other in the $(k-1)$-dimensional plane through $p'_1, p'_2, \cdots, p'_{k-1}$ and the center of the $S_{k-2,r}$ containing these $k-1$ points. Moreover

$$p'_k p^I_{k+1} \neq p'_k p^{II}_{k+1}$$

since p'_k does not lie on the $S_{k-2,r}$ containing $p'_1, p'_2, \cdots, p'_{k-1}$ and hence is not on the $(k-1)$-dimensional plane through $p'_1, p'_2, \cdots, p'_{k-1}$ and the center of the $S_{k-2,r}$.

Now we have

$$\begin{aligned} p'_1, p'_2, \cdots, p'_{k-1}, p^I_{k+1} &\approx p_1, p_2, \cdots, p_{k-1}, p_{k+1} \\ &\approx p'_1, p'_2, \cdots, p'_{k-1}, p^{II}_{k+1}. \end{aligned}$$

The points $p'_1, p'_2, \cdots, p'_{k-1}, p'_k, p^I_{k+1}, p^{II}_{k+1}$ all lie on the $S_{k-1,r}$. We wish to prove there exists a point p'_{k+1} such that $p'_1, p'_2, \cdots, p'_{k-1}, p'_k, p'_{k+1}$ are on the $S_{k,r}$ and

$$p'_1, p'_2, \cdots, p'_{k-1}, p'_k, p'_{k+1} \approx p_1, p_2, \cdots, p_{k-1}, p_k, p_{k+1}.$$

Consider the function $\Delta(x)$ defined as follows:

$$\Delta(x) = \begin{vmatrix} 1 & \cos\alpha_{12} & \cdots \cos\alpha_{1,k} & \cos\alpha_{1,k+1} \\ \cos\alpha_{21} & 1 & \cdots \cos\alpha_{2,k} & \cos\alpha_{2,k+1} \\ \cdot & \cdot & \cdots \quad \cdot & \cdot \\ \cos\alpha_{k,1} & \cos\alpha_{k,2} & \cdots \quad 1 & \cos x \\ \cos\alpha_{k+1,1} & \cos\alpha_{k+1,2} & \cdots \cos x & 1 \end{vmatrix} = A\cos^2 x + B\cos x + C;$$

where $A = -\Delta(p_1, p_2, \cdots, p_{k-1})$ which is negative, by (1). By Theorem II_{k-1}, $\Delta(x) = 0$ has the two roots $p'_k p^I_{k+1}/r$ and $p'_k p^{II}_{k+1}/r$, which we have shown to be unequal. Hence these are the only two roots in the interval $0 < x \leq \pi$. We may assume the labelling so that $p'_k p^I_{k+1}/r$ is less than $p'_k p^{II}_{k+1}/r$. Considering $\Delta(x)$ as a function of $\cos x$, we obtain a parabolic curve concave downward and crossing the $(\cos x)$-axis at the points

$$\cos(p'_k p^I_{k+1}/r) \quad \text{and} \quad \cos(p'_k p^{II}_{k+1}/r).$$

Since $\Delta(p_1, p_2, \cdots, p_{k+1})$ is positive, we have

$$\begin{aligned} \cos(p'_k p^{II}_{k+1}/r) &< \cos(p_k p_{k+1}/r) < \cos(p'_k p^I_{k+1}/r); \text{ that is} \\ p'_k p^I_{k+1}/r &< p_k p_{k+1}/r < p'_k p^{II}_{k+1}/r; \text{ or} \\ p'_k p^I_{k+1} &< p_k p_{k+1} < p'_k p^{II}_{k+1}. \end{aligned}$$

Consider the locus of all points p_x of the $S_{k,r}$ such that

$$p'_1 p_x = p_1 p_{k+1}; \quad p'_2 p_x = p_2 p_{k+1}; \quad \cdots; \quad p'_{k-1} p_x = p_{k-1} p_{k+1}.$$

This locus cuts the $S_{k-1,r}$ containing $p'_1, p'_2, \cdots, p'_{k-1}, p'_k, p^I_{k+1}, p^{II}_{k+1}$ at p^I_{k+1} and p^{II}_{k+1}. The function $p'_k p_x$ is a continuous function which takes on the values $p'_k p^I_{k+1}$ and $p'_k p^{II}_{k+1}$ and all values between these two values. Hence for some p_x, we have $p'_k p_x = p_k p_{k+1}$. Denote such a point by p'_{k+1}. Then we have

$$p_1, p_2, \cdots, p_{k-1}, p_k, p_{k+1} \approx p'_1, p'_2, \cdots, p'_{k-1}, p'_k, p'_{k+1}$$

and the theorem is proved.

THEOREM II'_k. *If $k + 2$ points $p_1, p_2, \cdots, p_{k+2}$ of a semi-metric space are such that each $k + 1$ points is congruent to $k + 1$ points of the $S_{k,r}$ and $\Delta(p_1, p_2, \cdots, p_{k+2}) = 0$, then the $k + 2$ points are congruent to $k + 2$ points of the $S_{k,r}$.*

If each $k + 1$ points is congruent to $k + 1$ points of the $S_{k-1,r}$ then the theorem is proved immediately by applying Theorem III_{k-1}. In this case the $k + 2$ points are congruent to $k + 2$ points of the $S_{k-1,r}$ and hence are congruent to $k + 2$ points of the $S_{k,r}$. We now have only to treat the case in which at least one $(k + 1)$-tuple is not congruent to $k + 1$ points of the $S_{k-1,r}$.

We may assume the labelling so that $p_1, p_2, \cdots, p_{k+1}$ are not congruent to $k + 1$ points of the $S_{k-1,r}$. By hypothesis these $k + 1$ points are congruent to $k + 1$ points of the $S_{k,r}$. Then, by Lemma 1, at least one k-tuple contained in these $k + 1$ points is congruent to k points of the $S_{k-1,r}$, but is not congruent to k points of the $S_{k-2,r}$. We shall assume the labelling so that $p_1, p_2, \cdots, p_k$ form such a k-tuple. Then

$$\Delta(p_1, p_2, \cdots, p_k) > 0.$$

By hypothesis there exist $k + 1$ points $p'_1, p'_2, \cdots, p'_k, p'_{k+1}$ of the $S_{k,r}$ and $k + 1$ points $\bar{p}_1, \bar{p}_2, \cdots, \bar{p}_k, \bar{p}_{k+2}$ of the $S_{k,r}$ such that

$$p_1, p_2, \cdots, p_k, p_{k+1} \approx p'_1, p'_2, \cdots, p'_k, p'_{k+1} \quad \text{and}$$
$$p_1, p_2, \cdots, p_k, p_{k+2} \approx \bar{p}_1, \bar{p}_2, \cdots, \bar{p}_k, \bar{p}_{k+2}.$$

We may make a congruent transformation of the $S_{k,r}$ into itself such that $\bar{p}_1$ goes into p'_1, $\bar{p}_2$ goes into $p'_2, \cdots, \bar{p}_k$ goes into p'_k. This transformation sends $\bar{p}_{k+2}$ into some point, p^*_{k+2}, of the $S_{k,r}$. The point p^*_{k+2} has its distances from the k points $p'_1, p'_2, \cdots, p'_k$ fixed. Since $p'_1, p'_2, \cdots, p'_k$ are not on an $S_{k-2,r}$, there are at most two possible positions for the point p^*_{k+2}. Two cases present themselves.

Case 1. *$p_1, p_2, \cdots, p_k, p_{k+2}$ are not congruent to $k + 1$ points of the $S_{k-1,r}$.* Then p^*_{k+2} is not on the $S_{k-1,r}$ containing $p'_1, p'_2, \cdots, p'_k$ and there are exactly two possible positions for the point p^*_{k+2}. We shall denote these two points by p^I_{k+2} and p^{II}_{k+2}. These two points are images of each other in the k-dimensional plane containing $p'_1, p'_2, \cdots, p'_k$ and the center of the sphere. The point p'_{k+1} is not on this plane since $p_1, p_2, \cdots, p_{k+1}$ are not congruent to $k + 1$ points of the $S_{k-1,r}$. Hence

$$p'_{k+1}p^I_{k+2} \neq p'_{k+1}p^{II}_{k+2}.$$

We define the expression $\Delta(x)$ as follows:

$$\Delta(x) = \begin{vmatrix} 1 & \cos\alpha_{12} & \cdots & \cos\alpha_{1,k+1} & \cos\alpha_{2,k+2} \\ \cos\alpha_{21} & 1 & \cdots & \cos\alpha_{2,k+1} & \cos\alpha_{2,k+2} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cos\alpha_{k+1,1} & \cos\alpha_{k+1,2} & \cdots & 1 & \cos x \\ \cos\alpha_{k+2,1} & \cos\alpha_{k+2,2} & \cdots & \cos x & 1 \end{vmatrix}.$$

$\Delta(x)$ is not identically zero since the coefficient of $\cos^2 x$, namely $-\Delta(p_1, p_2, \cdots, p_k)$ does not vanish. $\Delta(x) = 0$ has the two unequal roots $p'_{k+1}p^I_{k+2}/r$ and $p'_kp^{II}_{k+2}/r$ by the necessity of the conditions of Theorem II_k. By hypothesis $\Delta(x) = 0$ has the root $p_{k+1}p_{k+2}/r$. But $\Delta(x) = 0$ is a quadratic in $\cos x$ and hence has only two roots in the interval $0 < x \leq \pi$. Hence, we have

$$p_{k+1}p_{k+2} = p'_{k+1}p^I_{k+2} \quad \text{or} \quad p_{k+1}p_{k+2} = p'_{k+1}p^{II}_{k+2}.$$

Therefore either

$$p_1, p_2, \cdots, p_{k+1}, p_{k+2} \approx p'_1, p'_2, \cdots, p'_{k+1}, p^I_{k+2} \quad \text{or}$$
$$p_1, p_2, \cdots, p_{k+1}, p_{k+2} \approx p'_1, p'_2, \cdots, p'_{k+1}, p_{k+2}$$

and the theorem is proved.

Case 2. $p_1, p_2, \cdots, p_k, p_{k+2}$ *are congruent to* $k+1$ *points of the* $S_{k-1,r}$. In this case the point p^*_{k+2} is on the $S_{k-1,r}$ containing $p'_1, p'_2, \cdots, p'_k$ and hence is unique. We shall denote this point by p'_{k+2}. Evaluating $\Delta(x)$ as defined above we now have

$$(1) \quad \Delta(x) = \frac{-\begin{vmatrix} 1 & \cos\alpha_{12} & \cdots & \cos\alpha_{1,k} & \cos\alpha_{1,k+1} \\ \cos\alpha_{21} & 1 & \cdots & \cos\alpha_{2,k} & \cos\alpha_{2,k+1} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cos\alpha_{k,1} & \cos\alpha_{k,2} & \cdots & 1 & \cos\alpha_{k,k+1} \\ \cos\alpha_{k+2,1} & \cos\alpha_{k+2,2} & \cdots & \cos\alpha_{k+2,k} & \cos x \end{vmatrix}^2}{\Delta(p_1, p_2, \cdots, p_k)}.$$

By the same reasoning as in the preceding case we see that $\Delta(x)$ is not identically zero and that $\Delta(x) = 0$ has only two roots in the interval $0 < x \leq \pi$. By the necessity of the conditions of Theorem II_k a root of $\Delta(x) = 0$ is $p'_{k+1}p'_{k+2}/r$. By hypothesis a root of $\Delta(x) = 0$ is $p_{k+1}p_{k+2}/r$. From (1) we see that these must be double roots. Hence

$$p_{k+1}p_{k+2} = p'_{k+1}p'_{k+2}.$$

Therefore

$$p_1, p_2, \cdots, p_{k+1}, p_{k+2} \approx p'_1, p'_2, \cdots, p'_{k+1}, p'_{k+2}$$

and the proof is completed.

THEOREM III'_k. *If* $k+3$ *points* $p_1, p_2, \cdots, p_{k+3}$ *are such that each* $k+2$ *points is congruent to* $k+2$ *points of the* $S_{k,r}$ *and* $\Delta(p_1, p_2, \cdots, p_{k+3}) = 0$, *then the* $k+3$ *points are congruent to* $k+3$ *points of the* $S_{k,r}$.

If each $k+2$ points is congruent to $k+2$ points of the $S_{k-1,r}$ then the theorem is trivial since the $S_{k-1,r}$ has the congruence order $k+2$. Hence we shall suppose that at least one $(k+2)$-tuple, say $p_1, p_2, \cdots, p_{k+2}$, is not congruent to $k+2$ points of the $S_{k-1,r}$. By Theorem II_k, $\Delta(p_1, p_2, \cdots, p_{k+2})$ equals zero. Then at least one $(k+1)$-tuple contained in these $k+2$ points is not congruent to $k+1$ points of the $S_{k-1,r}$; for, by Theorem III_{k-1}, if each $(k+1)$-tuple is congruent to $k+1$ points of the $S_{k-1,r}$ and $\Delta(p_1, p_2, \cdots, p_{k+2})$ equals zero, the $k+2$ points are congruent to $k+2$ points of the $S_{k-1,r}$. We may assume the labelling so that $p_1, p_2, \cdots, p_{k+1}$ are not congruent to $k+1$ points of the $S_{k-1,r}$. But $p_1, p_2, \cdots, p_{k+1}$ are congruent to $k+1$ points of the $S_{k,r}$ since each $k+2$ points is congruent to $k+2$ points of the $S_{k,r}$. Hence,

$$\Delta(p_1, p_2, \cdots, p_{k+1}) > 0.$$

By hypothesis there exist $k+2$ points $p'_1, p'_2, \cdots, p'_{k+1}, p'_{k+2}$ of the $S_{k,r}$ and $k+2$ points $\bar{p}_1, \bar{p}_2, \cdots, \bar{p}_{k+1}, \bar{p}_{k+3}$ of the $S_{k,r}$ such that

$$p_1, p_2, \cdots, p_{k+1}, p_{k+2} \approx p'_1, p'_2, \cdots, p'_{k+1}, p'_{k+2} \quad \text{and}$$
$$p_1, p_2, \cdots, p_{k+1}, p_{k+3} \approx \bar{p}_1, \bar{p}_2, \cdots, \bar{p}_{k+1}, \bar{p}_{k+3}.$$

We may make a congruent transformation of the $S_{k,r}$ into itself so that $\bar{p}_1$ goes into p'_1, $\bar{p}_2$ goes into $p'_2, \cdots, \bar{p}_{k+1}$ goes into p'_{k+1}. This transformation sends $\bar{p}_{k+3}$ into some point, say p'_{k+3}. The point p'_{k+3} has its distances from the $k+1$ points $p'_1, p'_2, \cdots, p'_{k+1}$ fixed. These $k+1$ points are not on an $S_{k-1,r}$ since $p_1, p_2, \cdots, p_{k+1}$ are not congruent to $k+1$ points of the $S_{k-1,r}$. Hence the point p'_{k+3} is unique. We shall show that

$$p_1, p_2, \cdots, p_{k+2}, p_{k+3} \approx p'_1, p'_2, \cdots, p'_{k+2}, p'_{k+3}.$$

In order to do this we define the expression $\Delta(x)$ as follows:

$$\Delta(x) = \begin{vmatrix} 1 & \cos\alpha_{12} & \cdots & \cos\alpha_{1,k+2} & \cos\alpha_{1,k+3} \\ \cos\alpha_{21} & 1 & \cdots & \cos\alpha_{2,k+2} & \cos\alpha_{2,k+3} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cos\alpha_{k+2,1} & \cos\alpha_{k+2,2} & \cdots & 1 & \cos x \\ \cos\alpha_{k+3,1} & \cos\alpha_{k+3,2} & \cdots & \cos x & 1 \end{vmatrix}.$$

$\Delta(x)$ is not identically zero, for the coefficient of $\cos^2 x$, namely $-\Delta(p_1, p_2, \cdots, p_{k+1})$ does not vanish. $\Delta(x) = 0$ is a quadratic equation in $\cos x$ and hence has only two roots in the interval $0 < x \leq \pi$. By the necessity of the conditions of Theorem III_k, a root of $\Delta(x) = 0$ is $p'_{k+2}p'_{k+3}/r$. By hypothesis a root of $\Delta(x) = 0$ is $p_{k+2}p_{k+3}/r$. Evaluating $\Delta(x)$ we have

$$\Delta(x) = \frac{-\begin{vmatrix} 1 & \cos\alpha_{12} & \cdots & \cos\alpha_{1,k+1} & \cos\alpha_{1,k+2} \\ \cos\alpha_{21} & 1 & \cdots & \cos\alpha_{2,k+1} & \cos\alpha_{2,k+2} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cos\alpha_{k+1,1} & \cos\alpha_{k+1,2} & \cdots & 1 & \cos\alpha_{k+1,k+2} \\ \cos\alpha_{k+3,1} & \cos\alpha_{k+3,2} & \cdots & \cos\alpha_{k+3,k+1} & \cos x \end{vmatrix}^2}{\Delta(p_1, p_2, \cdots, p_{k+1})}.$$

Hence $p'_{k+2}p'_{k+3}/r$ and $p_{k+2}p_{k+3}/r$ are double roots of $\Delta(x) = 0$. Therefore

$$p_{k+2}p_{k+3} = p'_{k+2}p'_{k+3}.$$

Hence

$$p_1, p_2, \cdots, p_{k+2}, p_{k+3} \approx p'_1, p'_2, \cdots, p'_{k+2}, p'_{k+3}$$

which was to be shown.

THEOREM IV_k. *If $k+3$ points $p_1, p_2, \cdots, p_{k+3}$ form a pseudo $(k+3)$-tuple, then $\Delta(p_1, p_2, \cdots, p_{k+3})$ is negative.*

If each $k+2$ points are congruent to $k+2$ points of the $S_{k-1,r}$ the $k+3$

points do not form a pseudo $(k+3)$-tuple, since the $S_{k-1,r}$ has the congruence order $k+2$. Hence at least one $(k+2)$-tuple is not congruent to $k+2$ points of the $S_{k-1,r}$. We may assume the labelling so that $p_1, p_2, \cdots, p_{k+2}$ are not congruent to $k+2$ points of the $S_{k-1,r}$. By Theorem II_k, $\Delta(p_1, p_2, \cdots, p_{k+2})$ equals zero. Then at least one $(k+1)$-tuple contained in these $k+2$ points is not congruent to $k+1$ points of the $S_{k-1,r}$; for, by Theorem III_{k-1}, if each $k+1$ points is congruent to $k+1$ points of the $S_{k-1,r}$, and the determinant Δ is zero, then the $k+2$ points are congruent to $k+2$ points of the $S_{k-1,r}$. We may assume the labelling so that $p_1, p_2, \cdots, p_{k+1}$ are not congruent to $k+1$ points of the $S_{k-1,r}$. But $p_1, p_2, \cdots, p_{k+1}$ are congruent to $k+1$ points of the $S_{k,r}$. Then

$$\Delta(p_1, p_2, \cdots, p_{k+1}) > 0.$$

Evaluating $\Delta(p_1, p_2, \cdots, p_{k+3})$ we have

$$\Delta(p_1, p_2, \cdots, p_{k+3}) = \frac{-\begin{vmatrix} 1 & \cos\alpha_{12} & \cdots & \cos\alpha_{1,k+1} & \cos\alpha_{1,k+2} \\ \cos\alpha_{21} & 1 & \cdots & \cos\alpha_{2,k+1} & \cos\alpha_{2,k+2} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cos\alpha_{k+1,1} & \cos\alpha_{k+1,2} & \cdots & 1 & \cos\alpha_{k+1,k+2} \\ \cos\alpha_{k+3,1} & \cos\alpha_{k+3,2} & \cdots & \cos\alpha_{k+3,k+1} & \cos\alpha_{k+3,k+2} \end{vmatrix}^2}{\Delta(p_1, p_2, \cdots, p_{k+1})}.$$

Hence $\Delta(p_1, p_2, \cdots, p_{k+3})$ is less than or equal to zero. But if $\Delta(p_1, p_2, \cdots, p_{k+3})$ equals zero, the $k+3$ points are congruent to $k+3$ points of the $S_{k,r}$ by Theorem III_k. Hence $\Delta(p_1, p_2, \cdots, p_{k+3})$ is negative.

Theorems I_n, II_n, and III_n characterize sets of points which are congruent to $n+1$, $n+2$, and $n+3$ points of the $S_{n,r}$, respectively. Since the $S_{n,r}$ has the congruence order $n+3$, *the necessary and sufficient condition that a set of points containing more than $n+3$ points be congruent to a subset of the $S_{n,r}$ is that each $n+3$ of the points satisfy the conditions stated in Theorem III_n.*

THE RICE INSTITUTE. THE INSTITUTE FOR ADVANCED STUDY.

NOTE. It might be observed that the conditions obtained in this paper which serve to characterize r-spheric sets may be obtained in somewhat different form by an application of the theorems obtained by Menger for the n-dimensional euclidean space. The point of view adopted in the present investigation, however, is to regard the case of the n-dimensional spherical space as *fundamental*, characterizing r-spheric sets quite independently of the results referred to above, in order to obtain the conditions characterizing the n-dimensional euclidean space as well as the n-dimensional space of constant negative curvature by applying the results of this paper.

A BOUNDARY VALUE PROBLEM FOR THE HEAT EQUATION.*

By F. G. Dressel.

1. *Introduction.* In the first part of this paper, we shall be concerned with properties of a certain class of solutions of the equation

$$\partial^2 u/\partial x^2 - \partial u/\partial y = 0. \tag{1.1}$$

These solutions appear in the form of Stieltjes integrals. Such integrals have been used by various authors † in the study of discontinuous boundary value problems for Laplace's equation. In the second part, we give a generalization of a classical boundary value problem for the heat equation. Instead of assigning continuous boundary values to a solution of (1.1), we require that an integral of the solution take on boundary values which are preassigned functions of limited variation.

The function

$$u(x,y) = \int_a^b U(x,y;\xi,h)\,dF(\xi) \tag{1.2}$$

where

$$U(x,y;\xi,h) = [4\pi(y-h)]^{-1/2} \cdot e^{-[(x-\xi)^2/4(y-h)]}, \qquad y > h, \tag{1.3}$$
$$= 0, \qquad y \leq h,$$

and $F(\xi)$ is of limited variation in the closed interval (a, b), is a solution of (1.1) for $y > h$. Moreover, $u(x,y)$ is regular everywhere except perhaps on the segment $y = h$, $a \leq x \leq b$. The truth of this statement is obvious for $y \neq h$. If x is outside the interval (a, b), then $\lim_{y \to h^+} u(x,y) = 0$ by a fundamental property of the Stieltjes integral.‡ For the same reason, $\lim_{y\to h^+} \partial u/\partial y = 0$. It remains to investigate the behavior of $u(x,y)$ when the point (x,y) approaches, from above, an interior or a boundary point of the segment.

2. *Boundary values of the integral.* The point (x,y) will be said to approach the point (x_0, h), $h < y$, in the parabolic sense if there exist constants N and α, $\alpha > 1$, such that

* Presented, in part, to the Society, March 26, 1932.

† See bibliography given in G. C. Evans, "Complements of potential theory II," *American Journal of Mathematics*, Vol. 55 (1933), p. 29.

‡ Evans, *loc. cit.*, p. 14.

$$(x-x_0)^2 < N(y-h)^a.$$

We have the following theorem:

THEOREM 1. *If $M(x,y)$ approaches the point $P(x_0,h)$ in the parabolic sense, the function $u(M)$ defined by (1.2) takes on the value $F'(x_0)$ at every interior point of the segment (a,b) at which this derivative exists. If P is a boundary point of the segment, the respective limits are $F'(a+0)/2$ and $F'(b-0)/2$, provided these one-sided derivatives exist.*

Suppose first that $a < x_0 < b$, and replace ξ by $t = \xi - x$ in (1.2):

$$u(M) = \int_{a-x}^{b-x} U(t,y;0,h)\,dF(x+t).$$

Let $\rho(x+t,x_0)$ be defined by

$$F(x+t) = F(x_0) + (x+t-x_0)F'(x_0) + (x+t-x_0)\cdot\rho(x+t,x_0).$$

Then

$$u(M) = F'(x_0)\int_{a-x}^{b-x} U(t,y;0,h)\,dt + \int_{a-x}^{b-x} U(t,y;0,h)\,d_t[(x+t-x_0)\cdot\rho(x+t,x_0)].$$

The limit of the first term on the right is $F'(x_0)$, as may be verified by changing the variable of integration from t to $z = t/2(y-h)^{1/2}$. If in the second integral, we cut out an ϵ-neighborhood of $t=0$, the integrals which remain have the limit zero, since the integrands are continuous, and approach zero with $(y-h)$ uniformly for all t in the intervals. It suffices then to study the behavior of the integral extended over the neighborhood of $t=0$. On integrating by parts, we have

$$\int_0^{\epsilon} U(t,y;0,h)\,d[(x+t-x_0)\cdot\rho(x+t,x_0)] = U(t,y;0,h)(x+t-x_0)\cdot\rho(x+t,x_0)\Big]_0^{\epsilon} - \int_0^{\epsilon}(x+t-x_0)\rho(x+t,x_0)\,dU(t,y;0,h).$$

On account of the hypothesis on the manner of approach, the integrated part has the limit zero as $y \to h^+$. The integral which remains is equivalent to

$$\tfrac{1}{2}\int_0^{\epsilon}(y-h)^{-1}\cdot t^2\cdot\rho(x+t,x_0)U(t,y;0,h)\,dt + \tfrac{1}{2}\int_0^{\epsilon}(y-h)^{-1}\cdot(x-x_0)\cdot t\cdot\rho(x+t,x_0)U(t,y;0,h)\,dt = J_1 + J_2.$$

Denoting by $\bar{\rho}(\delta)$ the least upper bound of $|\rho(x+t, x_0)|$ in a circle of radius $\delta < \epsilon$ about the point $x = x_0$, $t = 0$, we have $\lim_{\delta \to 0} \bar{\rho}(\delta) = 0$.

$$|J_1| \leq 2 \cdot (\pi)^{-\frac{1}{2}} \cdot \bar{\rho} \cdot \int_0^{\epsilon/2(y-h)^{1/2}} z^2 \cdot e^{-z^2} dz = O[\bar{\rho}(\epsilon)]$$

J_2 approaches zero also since

$$\begin{aligned} |J_2| &\leqq (\bar{\rho}/2)\, |x - x_0| \cdot \int_0^{\epsilon} (y-h)^{-1} \cdot t \cdot U(t, y; 0, h)\, dt \\ &= -\bar{\rho} \cdot |x - x_0| \cdot U(t, y; 0, h) \Big]_0^{\epsilon} \\ &\leq \bar{\rho}(4\pi)^{-\frac{1}{2}} \cdot (y-h)^{(a-1)/2} \cdot (N)^{\frac{1}{2}} \{1 - e^{-[\epsilon^2/4(y-h)]}\} \\ &= O[\bar{\rho}(\epsilon)\, (y-h)^{(a-1)/2}]. \end{aligned}$$

This completes the proof if x_0 is an interior point of (a, b). If x_0 is a or b, we need only notice that under a parabolic approach, $|x - x_0|/(y-h)$ approaches zero as M approaches P, and then slight modifications of the preceding analysis give the results stated in the theorem.

For all $y > h$, $u(x, y)$ is continuous in x, and we may integrate it, thus defining the function

$$F(x, y) = \int_a^x u(t, y)\, dt, \qquad a \leq x \leq b, \quad h < y.$$

In order to study the behavior of $F(x, y)$ as y approaches h, evaluate $u(t, y)$, given by (1.2), by parts, and substitute in $F(x, y)$. We obtain

$$\begin{aligned} F(x, y) &= F(b) \int_a^x U(t, y; b, h)\, dt \\ &\quad - F(a) \int_a^x U(t, y; a, h)\, dt - \int_a^x dt \cdot \int_a^b F(\xi)\, d\, U(t, y; \xi, h) \\ &= I_1 + I_2 + I_3. \end{aligned}$$

We do not change the value of $U(x, y; \xi, h)$ if we interchange x and ξ; hence, by Theorem 1, the limit of I_1 is $F(b)/2$ or zero according as $x = b$ or $x < b$. Similarly, I_2 has the limit $-F(a)/2$ or zero according as $x > a$ or $x = a$. We may reverse the order of integration in I_3, and since $\partial U/\partial \xi = -\partial U/\partial t$, we have

$$I_3 = \int_a^b F(\xi) U(x, y; \xi, h)\, d\xi - \int_a^b F(\xi) U(a, y; \xi, h)\, d\xi.$$

These integrals are of a familiar type; it is known * that

* Goursat, *Cours d'Analyse Mathématique*, t. 3, p. 308.

$$\lim_{y\to h^+}\int_a^b F(\xi)U(x,y;\xi,h)\,d\xi = \frac{F(x+0)+F(x-0)}{2}, \qquad a<x<b,$$
$$= \frac{F(b-0)}{2}, \qquad x=b,$$
$$= \frac{F(a+0)}{2}, \qquad x=a.$$

We may accordingly state the theorem:

THEOREM 2. *If $u(x,y)$ is given by (1.2), the function*

$$F(x,y) = \int_a^x u(t,y)\,dt$$

has a limit as y approaches h^+. This limit is

$$\frac{F(x+0)+F(x-0)}{2} - \frac{F(a+0)+F(a)}{2},$$

where $F(x+0)=F(x)$ if $x=b$, and $F(x-0)=F(x)$ if $x=a$.

A function $u(x,y)$ of the type (1.2) will be said to belong to the class D if

$$\lim_{y\to h^+}[F(x_2,y)-F(x_1,y)]=0$$

for every set x_1, x_2 such that $a<x_1\leq x_2<b$. Making use of Theorem 2 and (1.2), one readily sees that all such $u(x,y)$ may be written in the form

$$u(x,y) = \alpha U(x,y;a,h) + \beta U(x,y;b,h)$$

where α and β are arbitrary constants.

3. *The integral analogous to the potential of a double layer.* Consider the new function

$$v(x,y) = \int_h^y V(x,y;\eta)\,dG(\eta), \qquad h\leq y\leq e, \tag{3.1}$$

where $G(\eta)$ is a function of limited variation in $h\leq\eta\leq e$, $\chi(\eta)$, $\chi'(\eta)$ are continuous, and

$$V(x,y;\eta) = -(2\partial/\partial x)U(x,y;\chi(\eta),\eta). \tag{3.2}$$

Since $V(x,y;y)=0$ if $x\neq\chi(y)$, (3.2) shows that, for all points (x,y) not on the curve $x=\chi(\eta)$, $v(x,y)$ is a solution of (1.1), regular in the band $h\leq y\leq e$. In considering the limit of $v(x,y)$ as (x,y) comes up to the curve $x=\chi(\eta)$, the following lemma will be useful:

LEMMA. *If*

$$I(x) = \int_0^c \frac{x-f(s)}{s^a}\cdot\phi(s,x)\,d[s\cdot t(s)], \qquad 0<c, \quad \alpha<2,$$

where

(a) *$f(s)$ is continuous, with a continuous derivative, $0 \leq s \leq c$, and* $f(0) = 0$;

(b) *$s \cdot t(s)$ is of limited variation, $0 \leq s \leq c$, and $t(s)$ is continuous at $s = 0$;*

(c) *$\phi(s, x)$ and $\partial\phi(s, x)/\partial s$ are continuous in s, $0 \leq s \leq c$, for all x, $0 < x \leq a$, and bounded $0 \leq x \leq a$; and $\phi(s, x)$ continuous in x for $s > 0$;*

(d) $\phi(0, x) = 0$, $x > 0$;

then

$$\lim_{x \to 0} I(x) = I(0).$$

The existence of $I(0)$ is insured by (b).* Write

$$I(x) = \int_\delta^c + \int_0^\delta \frac{x - f(s)}{s^a} \phi(s, x) d[s \cdot t(s)]$$
$$= I_1(x, \delta) + I_2(x, \delta).$$

From (a) and (c) the integrand in $I_1(x, \delta)$ converges uniformly to $f(s) \cdot \phi(s, 0)/s^a$ for $\delta > 0$, hence

$$\lim_{x \to 0} I_1(x, \delta) = I_1(0, \delta). \tag{1}$$

Evaluate $I_2(x, \delta)$ by parts:

$$I_2(x, \delta) = \delta^{2-a} \cdot [x - f(\delta)] \frac{\phi(\delta, x)}{\delta} t(\delta) - \int_0^\delta s \cdot t(s) d\left[\frac{x - f(s)}{s^a} \phi(s, x)\right].$$

Applying the mean value theorem to the fraction in the integrated part, we conclude from (c) and (d) that δ can be so chosen that for a given $\epsilon > 0$, the integrated part is in absolute value $< \epsilon/6$ for all x, $0 \leq x \leq a$. The integral which remains may be written

$$\int_0^\delta t(s) \left\{ \frac{\phi(s, x)}{s} \left[\frac{-ax}{s^{a-1}} - s^{2-a} \cdot f'(s) + a \cdot s^{2-a} \cdot \frac{f(s)}{s}\right] + \frac{x - f(s)}{s^{a-1}} \cdot \frac{\partial}{\partial s} \phi(s, x) \right\} ds.$$

Regarding these as four separate integrals, each converges to zero with δ, uniformly for all x in $0 \leq x \leq a$. Hence the absolute value of this integral is made $< \epsilon/6$ by taking δ suitably small, the inequality holding uniformly in x. Hence $| I_2(x, \delta) | < \epsilon/3$. That is, for all x in $0 \leq x \leq a$,

$$| I(x) - I_1(x, \delta) | < \epsilon/3. \tag{2}$$

---

* Evans, "Complements of potential theory I," *American Journal of Mathematics*, Vol. 54 (1932), p. 222.

Moreover

(3) $$| I(0) - I_1(0,\delta) | < \epsilon/3.$$

But $$| I(x) - I(0) | \leq | I_1(x,\delta) - I_1(0,\delta) | \\ + | I(x) - I_1(x,\delta) | + | I_1(0,\delta) - I(0) |.$$

From (2) and (3), δ may be chosen so small that each of the last two terms is $< \epsilon/3$ for all x in $0 \leq x \leq a$. With δ so chosen we take x small enough so that the first on the right is, by (1), $< \epsilon/3$. For x so chosen

$$| I(x) - I(0) | < \epsilon,$$

and this is what we wished to show.

We shall now prove that, if $G(\eta)$ has a derivative at $\eta = y$, then $v(x,y)$, for a fixed y, approaches a limit as x approaches $\chi(y)$, $x > \chi(y)$ and as x approaches $\chi(y)$, $x < \chi(y)$. The results are stated in the theorem:

THEOREM 3. *If $G(\eta)$ has a derivative at $\eta = y$, then for a fixed y*

$$\lim_{x \to \chi(y)} v(x,y) = \pm G'(y) + v(\chi(y), y),$$

the + or — being chosen according as x approaches the curve $x = \chi(\eta)$ from the right or from the left.

Write $G(\eta) = -(y - \eta)G'(y) - (y - \eta) \cdot t(\eta, y) + G(y)$

then $$\lim_{\eta \to y} t(\eta, y) = 0.$$

With the above substitution in $v(x,y)$ there results

$$v(x,y) = G'(y) \int_h^y V(x,y;\eta)d\eta - \int_h^y V(x,y;\eta)d[(y-\eta) \cdot t(\eta,y)].$$

The first term on the right has the limit as x approaches $\chi(y)$ *

$$\pm G'(y) + G'(y) \int_h^y V(\chi(y), y; \eta)d\eta.$$

Our lemma is seen to apply to the second integral of $v(x,y)$, hence the limit of the integral is equal to the integral of the limit. Combining these results the theorem follows.

In the band $h \leq y \leq e$ the function $v(x,y)$ defined by (3.1) is continuous for all points (x,y) off the curve

$$C: \quad x = \chi(\eta), \qquad h \leq \eta \leq e.$$

---

* Goursat, *loc. cit.*, p. 308.

We can therefore form its line integral along curves that do not have contact with C. In particular if we consider a parallel displacement of C in the x-direction, we may integrate $v(x, y)$ along such a curve, forming the function:

$$G(x,y) = \int_h^y v(\chi(t) + \lambda, t)\, dt \tag{3.3}$$

where λ represents the distance from the curve C to the displaced curve, and is equal to $[x - \chi(y)]$.

For $\lambda \neq 0$, $G(x, y)$ is seen to be a continuous function of the point (x, y). We wish then to examine the limiting value $G(x, y)$ takes on, if we fix y and let x approach the curve C through values of $x < \chi(y)$ and also for values of $x > \chi(y)$. To this end replace the integrand in (3.3) by its value given by (3.1); for $\lambda \neq 0$, in the resulting function we may change the order of integration. Considering that this has been done, on adding and subtracting the function

$$H(x,y) = \int_h^y dG(\eta) \cdot \int_\eta^y 2\chi'(t) U(\chi(t) + \lambda, t; \chi(\eta), \eta)\, dt, \tag{3.4}$$

there results

$$\begin{aligned} G(x,y) = \int_h^y dG(\eta) \cdot \int_\eta^y [&V(\chi(t) + \lambda, t; \eta) \\ &- 2\chi'(t) U(\chi(t) + \lambda, t; \chi(\eta), \eta)]\, dt + H(x,y). \end{aligned} \tag{3.5}$$

Without loss of generality we assume $G(h) = 0$. Since

$$\begin{aligned} &(\partial/\partial\eta)[V(\chi(t) + \lambda, t; \eta) - 2\chi'(t) \cdot U(\chi(t) + \lambda, t; \chi(\eta), \eta)] \\ &\quad = -(\partial/\partial t)[V(\chi(t) + \lambda, t; \eta) - 2\chi'(\eta) \cdot U(\chi(t) + \lambda, t; \chi(\eta), \eta)], \end{aligned}$$

and since

$$[V(\chi(\eta) + \lambda, \eta; \eta) - 2\chi'(\eta) \cdot U(\chi(\eta) + \lambda, \eta; \chi(\eta), \eta)] = 0$$

then on integrating by parts the first term on the right side of (3.5) we get

$$\begin{aligned} G(x,y) = \int_h^y [&V(\chi(y) + \lambda, y; \eta) \\ &- 2\chi'(\eta) U(\chi(y) + \lambda, y; \chi(\eta), \eta)] G(\eta)\, d\eta + H(x,y). \end{aligned} \tag{3.6}$$

Again using the limit theorem on Stieltjes integrals, we have

$$\lim_{\lambda \to 0} H(x,y) = H(\chi(y), y).$$

The first integral is more troublesome. In it we may assume that $G(\eta)$ is non-decreasing. For a $\delta > 0$, since $G(\eta)$ is a function of limited variation, we can select ϵ so that for $y - \epsilon \leq \eta \leq y$, we have $[G(y - 0) - G(\eta)] < \delta$. Write the first integral on the right in (3.6) in the form

$$(3.7)\quad \int_h^{y-\epsilon} + \int_{y-\epsilon}^{y} [V(\chi(y)+\lambda, y;\eta)$$
$$-2\chi'(\eta)U(\chi(y)+\lambda, y;\chi(\eta),\eta)]\,[G(\eta)-G(y-0)]d\eta$$
$$+\,G(y-0)\int_h^y V(\chi(y)+\lambda, y;\eta)d\eta$$
$$-\,G(y-0)\int_h^y 2\chi'(\eta)U(\chi(y)+\lambda, y;\chi(\eta),\eta)d\eta.$$

The first integral above has a continuous integrand, the fourth integral is seen to be uniformly convergent, hence for these the limits of the integrals as λ approaches zero are equal to the integrals of the respective limits. The third integral, as we have just seen in proving Theorem 3, has the limiting value

$$\pm\, G(y-0) + G(y-0)\int_h^y V(\chi(y), y;\eta)d\eta.$$

On applying the Second Law of the Mean * to the second integral, there results

$$[G(y-\epsilon+0)-G(y-0)]\cdot\int_{y-\epsilon}^{\xi} [V(\chi(y)+\lambda, y;\eta)$$
$$-2\chi'(\eta)\cdot U(\chi(y)+\lambda, y;\chi(\eta),\eta)]d\eta, \qquad y-\epsilon<\xi<y.$$

Making the substitution $[\chi(y)+\lambda-\chi(\eta)]/2(y-\eta)^{1/2}=z$ in the integral we see the above term in absolute value can be made less than

$$\delta\cdot 2\cdot(\pi)^{-1/2}\cdot\int_{-\infty}^{+\infty} e^{-z^2}\,dz.$$

Since δ can be taken arbitrarily small, we conclude that the second integral on the right in (3.7) approaches zero with ϵ uniformly in λ. These results give the important theorem:

THEOREM 4. *If we write*

$$\bar{G}(\eta) = G(\eta) - G(h)$$

then, for a fixed y, we have

$$\lim_{x\to\chi(y)} G(x,y) = \pm\,\bar{G}(y-0) + \int_h^y [V(\chi(y), y;\eta)$$
$$-2\chi'(\eta)U(\chi(y), y;\chi(\eta),\eta)]\,\bar{G}(\eta)d\eta + H(\chi(y), y),$$

the $+$ or $-$ sign being taken according as the approach is through values of $x>\chi(y)$, or $x<\chi(y)$.

---

* Evans, *The Logarithmic Potential*, p. 17.

Holmgren's theorem * is the particular case of Theorem 4 in which $G(y)$ has a derivative satisfying a Hölder condition of order $> \frac{1}{2}$.

The function

$$t(x,y) = \int_h^y 2\chi'(\eta) U(x,y;\chi(\eta),\eta)\, dG(\eta)$$

is seen to be a solution of (1.1) which is regular in the band $h \leq y \leq e$, for all points (x,y) off the curve C. If we form its line integral along the curve $x = \chi(\xi) + \lambda$, we have for $\lambda \neq 0$

$$(3.8) \qquad \int_h^y t(\chi(\xi)+\lambda,\xi)\, d\xi = \int_h^y dG(\eta) \cdot \int_\eta^y 2\chi'(\eta) U(\chi(\xi)+\lambda,\xi;\chi(\eta),\eta)\, d\xi,$$

the form on the right being obtained by replacing $t(x,y)$ by its value given above and then reversing the order of integration. As for the similar function $H(x,y)$ we have for a fixed y

$$(3.9) \qquad \lim_{\lambda\to 0} \int_h^y t(\chi(\xi)+\lambda,\xi)\, d\xi = \int_h^y dG(\eta) \cdot \int_\eta^y 2\chi'(\eta) \cdot U(\chi(\xi),\xi;\chi(\eta),\eta)\, d\xi.$$

If we define

$$(3.10) \qquad w(x,y) = v(x,y) - t(x,y),$$

then making use of (3.9) we may state the following corollary to Theorem 4:

COROLLARY. *If $\chi(\eta)$ has a continuous second derivative, and*

$$\bar{G}(x,y) = \int_h^y w(\chi(t)+\lambda,t)\, dt$$

then, for a fixed y,

$$\lim_{x\to\chi(y)\pm} \bar{G}(x,y) = \pm\, \bar{G}(y-0) + \int_h^y [V(\chi(y),y;\eta) - 2\chi'(\eta) U(\chi(y),y;\chi(\eta),\eta)]\bar{G}(\eta)\, d\eta + \int_h^y \{(\partial/\partial\eta) \int_\eta^y 2[\chi'(\eta)-\chi'(t)]\, U(\chi(t),t;\chi(\eta),\eta)\, dt\}\bar{G}(\eta)\, d\eta.$$

4. *A generalized boundary value problem.* Suppose D is a finite domain bounded by the characteristics whose ordinates are h and e $(h < e)$, and the arcs

* Goursat, *loc. cit.*, p. 306.

$$x = \chi_i(y), \qquad h \leq y \leq e, \qquad (i = 1, 2),$$

where $\chi_1 < \chi_2$, and the functions $\chi_i, \chi'_i, \chi''_i$ are continuous. Segments of the curves $y = h + \delta$, $x = \chi_i(y) + \lambda_i$ where $\delta > 0$, $(-1)^{i+1} \cdot \lambda_i > 0$ will be called δ- and λ_i-displacements respectively. We propose to demonstrate the existence of a function $l(x, y)$ which will be a solution of (1.1) within D; and whose line integral along a δ-displacement lying within D will have as δ approaches zero the limiting value $[B(x_2) - B(x_1)]$, where x_2 and x_1 $(x_2 \geq x_1)$ are the abscissas of the ends of the δ-displacement, and $B(\xi)$ is a given function of limited variation with regular discontinuities $\chi_1(h) \leq \xi \leq \chi_2(h)$, ($\delta$-displacements having abscissas $\chi_1(h)$ or $\chi_2(h)$ as end points are excluded); also the line integral of $l(x, y)$ along a λ_i-displacement will have as λ_i approaches zero the limit $[G_i(y_2) - G_i(y_1)]$, y_2 and y_1 $(y_2 \geq y_1)$ are the ordinates of the ends of the λ_i-displacement, and the $G_1(\eta)$, $G_2(\eta)$ are preassigned functions of limited variation continuous from the left for $h \leq \eta \leq e$.

In $V(x, y; \eta)$ given by (3.2), replace $\chi(\eta)$ by $\chi_i(\eta)$, and denote the resulting function by $V_i(x, y; \eta)$; then consider the function

$$\begin{aligned}(4.1) \qquad l(x, y) = & \int_h^y [V_1(x, y; \eta) - 2\chi'_1(\eta) U(x, y; \chi_1(\eta), \eta)] \, dF_1(\eta) \\ & + \int_h^y [V_2(x, y; \eta) - 2\chi'_2(\eta) U(x, y; \chi_2(\eta), \eta)] \, dF_2(\eta) \\ & + \int_a^b U(x, y; \xi, h) dB(\xi); \qquad a = \chi_1(h), \ b = \chi_2(h) \\ = & \ w_1(x, y) + w_2(x, y) + u(x, y).\end{aligned}$$

We see that $l(x, y)$ in the domain D satisfies (1.1) if $F_i(\eta)$ are functions of limited variation, moreover by Theorem 2 the line integral of $u(x, y)$ along a δ-displacement will have $[B(x_2) - B(x_1)]$ as a limit as δ approaches zero. Along a permissible δ-displacement $| x - \chi_i(\eta) | > 0$, therefore the integrands in $w_i(x, y)$ are continuous and approach zero as y approaches h, hence

$$(4.2) \qquad \lim_{y \to h^+} \int_{x_1}^{x_2} l(t, y) dt = B(x_2) - B(x_1), \qquad a < x_1 \leq x_2 < b.$$

Thus we have yet to show the existence of functions $F_i(\eta)$ which are of limited variation, continuous from the left, and which satisfy the conditions

$$(4.3) \qquad \lim_{\lambda_i \to 0} \int_{y_1}^{y_2} l(\chi_i(t) + \lambda_i, t) dt = G_i(y_2) - G_i(y_1).$$

By Lebesgue's theorem on limits of integrals, for a fixed y,

$$(4.4) \qquad \lim_{\lambda_i \to 0} \int_h^y u(\chi_i(t) + \lambda_i, t) dt = \int_h^y u(\chi_i(t), t) dt.$$

The function on the right is evidently absolutely continuous. Thus from (4.4) and the corollary to Theorem 4 the pair of equations (4.3), taking $y_1 = h$ and $y_2 = y$, can be put in the form *

$$f_i(y) = g_i(y) + \int_h^y K_{ij}(y, \eta) f_j(\eta)\, d\eta, \qquad (j = 1, 2), \tag{4.5}$$

where the $g_i(y)$ are known functions of limited variation continuous from the left

$$g_i(y) = (-1)^{i+1} \cdot [G_i(y) - G_i(h) - \int_h^y u(\chi_i(t), t)\, dt]$$

and

$$f_i(y) = F_i(y) - F_i(h) \tag{4.6}$$

$$K_{ij}(y, \eta) = (-1)^i [V_j(\chi_i(y), y; \eta) - 2\chi'_j(\eta) U(\chi_i(y), y; \chi_j(\eta), \eta) + (\partial/\partial\eta)\{\int_\eta^y 2[\chi'_j(\eta) - \chi'_i(t)] \cdot U(\chi_i(t), t; \chi_j(\eta), \eta)\, dt\}]. \tag{4.7}$$

Each kernel $K_{ij}(y, \eta)$ is of the form of a continuous function of y and η divided by $(y - \eta)^{1/2}$, hence one iteration of the system (4.5) will produce a system having continuous kernels. Such a system will have a unique solution continuous from the left, hence also will the system (4.5). We shall now show that the solutions are of limited variation.

5. *The solutions in terms of Stieltjes integrals.* Write the solutions of (4.5) in the form †

$$f_i(y) = g_i(y) + \sum_{\beta=1}^{\infty} \int_h^y K_{ij}^{[\beta]}(y, \eta) g_j(\eta)\, d\eta \tag{5.1}$$

where

$$K_{ij}^{[1]}(y, \eta) = K_{ij}(y, \eta) \tag{5.2}$$

$$K_{ij}^{[\beta]}(y, \eta) = \int_\eta^y K_{ik}^{[\alpha]}(y, \xi) K_{kj}^{[\beta-\alpha]}(\xi, \eta)\, d\xi,$$

$$(k = 1, 2), \quad (\alpha = 1, \cdots, \beta - 1).$$

Remembering that $g_i(h) = 0$ the equations (5.1) may be written

$$f_i(y) = g_i(y) + \sum_{\beta=1}^{\infty} \int_h^y dg_j(\eta) \int_\eta^y K_{ij}^{[\beta]}(y, \xi)\, d\xi. \tag{5.3}$$

The functions

$$\int_\eta^y K_{ij}^{[\beta]}(y, \xi)\, d\xi$$

* The usual convention of summation with respect to an index repeated in the same term is used throughout the rest of the paper.

† V. Volterra, *Leçons sur les Équations Intégrales*, p. 71.

need to be put into another form; since

$$\int_\eta^y [V_j(\chi_i(y), y; \xi) - 2\chi'_j(\xi) U(\chi_i(y), y; \chi_j(\xi), \xi)] d\xi$$
$$= \int_\eta^y [V_j(\chi_i(\xi), \xi; \eta) - 2\chi'_i(\xi) U(\chi_i(\xi), \xi; \chi_j(\eta), \eta)] d\xi$$

we have, after integrating the last terms in (4.7) and combining it with the right member of the preceding equation, that for $\beta = 1$

$$\int_\eta^y K_{ij}^{[1]}(y, \xi) d\xi = \int_\eta^y T_{ij}^{[1]}(\xi, \eta) d\xi$$

where

$$T_{ij}^{[1]}(\xi, \eta) = (-1)^i [V_j(\chi_i(\xi), \xi; \eta) - 2\chi'_j(\eta) U(\chi_i(\xi), \xi; \chi_j(\eta), \eta)].$$

Making use of the above equations and induction it is readily shown that

$$\int_\eta^y K_{ij}^{[\beta]}(y, \xi) d\xi = \int_\eta^y T_{ij}^{[\beta]}(\xi, \eta) d\xi$$

where

$$T_{ij}^{[\beta]}(\xi, \eta) = \int_\eta^\xi T_{ik}^{[\alpha]}(\xi, t) T_{kj}^{[\beta-\alpha]}(t, \eta) dt, \qquad (\beta = 2, 3, \cdots). \tag{5.4}$$

With the above substitution the solutions (5.3) become

$$f_i(y) = g_i(y) + \sum_{\beta=1}^{\infty} \int_h^y dg_j(\eta) \int_\eta^y T_{ij}^{[\beta]}(\xi, \eta) d\xi. \tag{5.5}$$

From (5.4) we see there is a finite number M such that

$$| T_{ij}^{[\beta]}(\xi, \eta) | < M, \qquad (\beta = 2, 3), \tag{5.6}$$

whence we easily deduce that

$$| T_{ij}^{[\beta]}(\xi, \eta) | < 2^{\gamma-1} \cdot M^\gamma \cdot [(e - h)^{\gamma-1}/(\gamma - 1)!], \qquad (\beta \geqq 2) \tag{5.7}$$

where γ equals $\beta/2$ or $(\beta - 1)/2$ according as β is an even or an odd integer.

For $\beta \geqq 2$, one finds on setting up the expression for the total variation of

$$\int_h^y dg_j(\eta) \int_\eta^y T_{ij}^{[\beta]}(\xi, \eta) d\xi \tag{5.8}$$

and using (5.6) and (5.7) that it can not exceed

$$2^{\gamma+1} \cdot M^\gamma \cdot [e - h)^\gamma/(\gamma - 1)!] \cdot R$$

where R is the larger of the total variations of $g_1(\eta)$ and $g_2(\eta)$. Thus we see that

$$\sum_{\beta=2}^{\infty}\int_h^v dg_i(\eta)\int_\eta^v T_{ij}^{[\beta]}(\xi,\eta)\,d\xi$$

is of limited variation since its total variation can not exceed the sum of the convergent series

$$\sum_{\gamma=1}^{\infty} 2^{\gamma+2}\cdot M^\gamma[(e-h)^\gamma/(\gamma-1)\,!\,]R.$$

The next lemma will show that the functions of the form (5.8) for $\beta=1$ are of limited variation, hence the right members of (5.5) are such functions.

LEMMA. *If $g(\eta)$ is of limited variation for $h\leq\eta\leq e$ then the function*

$$\Phi(y)=\int_h^v dg(\eta)\int_\eta^v T_{ij}^{[1]}(\xi,\eta)\,d\xi$$

is of limited variation.

Define

$$\begin{aligned}\alpha_{ij}(y,\eta) &= \int_\eta^v T_{ij}(\xi,\eta)\,d\xi, && \eta<y,\\ &=0, && \eta\geq y,\end{aligned}$$

then $\Phi(y)$ becomes

$$\int_h^e \alpha_{ij}(y,\eta)\,dg(\eta).$$

It is evident that $\alpha_{ij}(y,\eta)$ is continuous in η and of uniformly limited variation in y, hence the above integral is of limited variation.* Thus we have established:

THEOREM 5. *For a given set of functions $g_i(y)$ of limited variation continuous from the left, and vanishing at h, there exists a unique set of functions $f_i(y)$ satisfying* (4.5) *and having those same properties.*

The equations (4.6) determine the functions $F_i(y)$ to within additive constants, but additive constants on the functions $F_i(y)$ do not affect the function $l(x,y)$ given by (4.1). Turning then to the $g_i(y)$, we see that once the function $u(x,y)$ in (4.1) is fixed the $g_i(y)$ are unique, the function $u(x,y)$, however, is unique to within an additive function of the class D. We have then the final theorem:

THEOREM 6. *In the class of functions* (4.1) *there exists a function $l(x,y)$, unique to within an additive function of class D, such that $l(x,y)$ is a solution of* (1.1) *and satisfies the conditions* (4.2) *and* (4.3).

DUKE UNIVERSITY.

---

* H. E. Bray, "Elementary properties of the Stieltjes Integral," *Annals of Mathematics*, vol. 20 (1919), p. 181.

ON ABSTRACT CLOSED SURFACES OF NEGATIVE CURVATURE.

By Monroe H. Martin.*

It is well known † that there exist abstract closed surfaces of an arbitrary genus $p > 1$ possessing constant negative curvature. Such a surface may be constructed by considering a Fuchsian group possessing the real axis of the complex plane as the principal circle of the group and a fundamental domain lying entirely above the real axis bounded by arcs of $4p (p > 1)$ circles orthogonal to the real axis. The transformations of this group carry the fundamental domain into congruent regions which cover the *entire* upper half-plane without lacunae. Furthermore, they leave the differential form

$$ds^2 = (dx^2 + dy^2)/y^2$$

invariant. Now if we consider congruent points as identical and do not insist that the *entire* upper half-plane, possessing this metric, permit an isometric mapping on some real surface possessing a continuously turning normal at all points of the surface, we may say that an abstract closed surface possessing constant negative curvature of genus $p > 1$ is obtained. On the other hand, suppose we require that the fundamental domain, together with the totality of all regions congruent to it, i. e. the *entire* upper half-plane be realized on some real surface in the above manner and attempt to construct an abstract closed surface by considering congruent points on this real surface as identical. From this point of view the construction is impossible for Hilbert ‡ has demonstrated that there exists no real surface possessing a continuously turning normal at all points of the surface which may be mapped isometrically on the *entire* upper half-plane possessing the above metric.

If one considers surfaces, not of constant negative curvature, but possessing negative curvature everywhere the following question naturally arises:

Do there exist surfaces possessing a continuousuly turning normal at all points of the surface and negative curvature everywhere, which permit an infinite discrete group of transformations into themselves such that when con-

---

* National Research Fellow.

† See, for instance, H. M. Morse, "Geodesics on closed surfaces." *Transactions of the American Mathematical Society*, Vol. 26 (1924), pp. 26-33.

‡ D. Hilbert, "Über Flächen von konstanter Gausscher Krümmung," *Transactions of the American Mathematical Society*, Vol. 2 (1901), pp. 87-99.

gruent points are considered as identical an abstract, closed surface of negative curvature everywhere and of genus $p \geqq 0$ *is obtained?*

As a partial answer to this question we shall show that the answer is in the affirmative and that the surfaces may even be analytic for $p > 1$. The answer to the question for $p = 0$ and $p = 1$ is still open. As a step in this direction there is obtained an abstract, closed surface of genus $p = 1$ on which the curvature is negative everywhere with the exception of two points of zero curvature. An example is also given of a non-orientable abstract, closed surface possessing negative curvature everywhere with the exception of four points of zero curvature.

Preparatory to constructing a surface S possessing the properties given above it is first necessary to consider in some detail the analytic surface defined by the equation

$$(1) \qquad f(x, y, z) = \cos x + \cos y + \cos z = 0.$$

The curvature K of this surface at any point (x, y, z) is most readily calculated from a formula of Gauss for the curvature of a surface when its equation is given in the form $f(x, y, z) = 0$ and turns out to be

$$K = -\frac{\cos^2 x + \cos^2 y + \cos x \cos y}{(\sin^2 x + \sin^2 y + \sin^2 z)^2}.$$

The curvature K may then readily be shown to be, because of the inequality

$$\cos^2 x + \cos^2 y \geqq 2 \mid \cos x \mid \ \mid \cos y \mid,$$

negative everywhere, except at the points where $\cos x$ and $\cos y$ vanish simultaneously, namely the points

$$(\pm \pi/2 + 2k\pi, \quad \pm \pi/2 + 2l\pi, \quad \pm \pi/2 + 2m\pi) \quad (k, l, m = 0, \pm 1, \cdots),$$

at which it is zero. A simple calculation shows that the above points are all umbilical points of our surface. The traces of the surface on the planes

$$A: x = k\pi, \quad B: y = l\pi, \quad C: z = m\pi \qquad (k, l, m = 0, \pm 1, \cdots)$$

have the equations

$$(2) \qquad a: \cos y + \cos z = (-1)^{k+1}, \qquad b: \cos z + \cos x = (-1)^{l+1}, \\ c: \cos x + \cos y = (-1)^{m+1},$$

respectively. Consider now the curves c. For a given value of m they form a set of convex, symmetrical and analytic ovals, invariant under the group of translations

$$\bar{x} = x + 2k\pi, \qquad \bar{y} = y + 2l\pi \qquad\qquad (k, l = 0, \pm 1, \cdots),$$

and lying in the plane $z = m\pi$. For even values of m the coördinates of the centers of symmetry of the ovals belonging to the set are given by

$$((2p+1)\pi,\ (2q+1)\pi,\ m\pi) \qquad (p, q = 0, \pm 1, \cdots)$$

and for odd values of m by

$$(2p\pi,\ 2q\pi,\ m\pi) \qquad (p, q = 0, \pm 1, \cdots).$$

The trace of the surface on the plane

$$z = m\pi + \theta\pi \qquad 0 \leqq \theta \leqq 1 \tag{3}$$

obviously varies from the set of ovals in the plane $z = m\pi$ for $\theta = 0$ to the set of ovals in the plane $z = (m+1)\pi$ for $\theta = 1$. Consequently if θ be imagined to increase monotonely from 0 to 1 the ovals comprising the trace of the surface on the plane (3) expand and at first possess no points in common until the value $\theta_0 = \frac{1}{2}$ of the parameter θ is reached. At this point double points occur and after this point the trace of the surface on the plane (3) shrinks down into the trace of the surface on the plane $z = (m+1)\pi$. Because of the symmetrical character of the equation (1) of our surface it is clear that the set of ovals in any one of the two remaining families of planes is congruent to the set of ovals contained in the family of planes C. For example, the set of ovals lying in the family of planes A may be sent into the set of ovals lying in the family of planes C by the rotation + translation

$$\bar{x} = -z + \pi, \quad \bar{y} = y + \pi, \quad \bar{z} = x - \pi$$

which represents a rotation of 90° about the line $x = \pi$, $z = 0$ followed by a translation for a distance π in the direction of the positive y-axis. For our purposes it is necessary to divide each of the three sets of ovals a, b, c lying respectively in the three families of planes A, B, C into two subsets. We divide a into two subsets a^+ and a^- according as k in (2) is odd or even, b into two subsets b^+ and b^- according as l in (2) is odd or even, and c into two subsets c^+ and c^- according as m in (2) is odd or even. It is then readily verified that the coördinates of the centers of symmetry of the ovals belonging to these six subsets arrange themselves as follows:

$$a^+ : (k\pi, 2p\pi, 2q\pi), \qquad b^+ : (2p\pi, l\pi, 2q\pi),$$
$$a^- : (k\pi, (2p+1)\pi, (2q+1)\pi); \quad b^- : ((2p+1)\pi, l\pi, (2q+1)\pi);$$
$$c^+ : (2p\pi, 2q\pi, m\pi),$$
$$c^- : ((2p+1)\pi, (2q+1)\pi, m\pi);$$
$$(p, q = 0, \pm 1, \cdots).$$

We next observe that the surface is taken into itself by the transformations of the linear group

$$(4)\quad \begin{array}{ll} \mathrm{I}: \begin{array}{l}\bar{x} = -y + (q_1 + p_1)\pi \\ \bar{y} = x + (q_1 - p_1)\pi\end{array}; & \mathrm{II}: \begin{array}{l}\bar{x} = -x + 2p_2\pi \\ \bar{y} = -y + 2q_2\pi\end{array}; \\ \mathrm{III}: \begin{array}{l}\bar{x} = y + (p_3 - q_3)\pi \\ \bar{y} = -x + (p_3 + q_3)\pi\end{array}; & \mathrm{IV}: \begin{array}{l}\bar{x} = x + 2p_4\pi \\ \bar{y} = y + 2q_4\pi\end{array}; \\ \multicolumn{2}{c}{\mathrm{V}: \bar{z} = z + 2m\pi,} \end{array}$$

where the p_i and q_i are any integers positive, negative, or zero subject to the restriction that the pairs p_i, q_i for which $i = 1, 3$ are formed from integers of the same parity and m is any integer whatsoever. Those transformations of the group (4) which are of type I obviously have as their geometrical representation a rotation in the positive sense about the line $x = p_1\pi$, $y = q_1\pi$ through an angle of $90°$; those of type II a rotation in the positive sense about the line $x = p_2\pi$, $y = q_2\pi$ through an angle of $180°$; those of type III a rotation in the positive sense about the line $x = p_3\pi$, $y = q_3\pi$ through an angle of $270°$. The transformations of the group (4) of the types IV and V are translations.

Let us designate the parallelopiped

$$(5)\qquad 0 < x < \pi, \quad 0 \leqq y \leqq \pi, \quad -\pi \leqq z < +\pi$$

as the fundamental parallelopiped and the portion of the surface (1) contained therein as the fundamental domain. From the geometrical interpretation of the transformations of the group (4) and the peculiar structure of the surface (1) it follows that the surface (1) may be seen as infinitely many copies of the fundamental domain distributed throughout the entire x, y, z-space by the transformations of the group (4). If one adopts the convention that all points in x, y, z-space which are obtained from points of the fundamental domain by the transformations of the group (4) are identical, an abstract, closed surface, possessing negative curvature everywhere with the exception of two points of zero curvature, is obtained. In order to determine the connectivity of this abstract, closed surface we first observe that the portion of the boundary of the fundamental domain comprised by the semi-circumference of an oval belonging to the set b^+ is carried into that portion of the boundary of the fundamental domain comprised by the semi-circumference of an oval belonging to the set a^+ by the rotation

$$\bar{x} = -y + 2\pi, \quad \bar{y} = x$$

about the line $x = \pi$, $y = \pi$ belonging to the group (4). Secondly, we observe that the portion of the boundary of the fundamental domain formed by two

quadrants of ovals belonging to the set b^- is carried into the portion of the boundary of the fundamental domain formed by two quadrants of ovals belonging to the set a^- by the rotation

$$\bar{x} = -y, \quad \bar{y} = x,$$

about the line $x = 0$, $y = 0$ belonging to the group (4). According to the above convention we consider as identical all points of the boundary of the fundamental domain which are obtained from one another by means of the transformations of the group (4). At this point we then have a surface homeomorphic with a finite cylinder possessing two boundaries which correspond to the two boundaries of the fundamental domain formed by two quadrants of ovals belonging to the set c^+. Finally, since these two quadrants are carried into one another by a translation

$$\bar{z} = z + 2\pi,$$

belonging to the group (4), the two boundaries of the cylinder are united to obtain a surface of genus $p = 1$.

If one takes as fundamental domain the portion of the surface contained in the cube

$$-\pi \leqq x < +\pi, \quad -\pi \leqq y < +\pi, \quad -\pi \leqq z < +\pi$$

and as the group of transformations of the surface into itself the subgroup of the group (4) formed by the transformations of the types IV and V in (4), an abstract, closed surface of genus $p = 3$ is obtained which possesses negative curvature everywhere with the exception of eight points of zero curvature.

We are now in a position to construct a surface S possessing negative curvature everywhere and admitting the group (4). In order to do this we note that the surface (1) divides the entire x, y, z-space into two regions according as $f(x, y, z) > 0$ or $f(x, y, z) < 0$. The first of these regions we shall speak of as the positive region and the second as the negative region. The centers of symmetry of the ovals belonging to the sets a^+, b^+, c^+ lie in positive region and the centers of symmetry of the ovals belonging to the sets a^-, b^-, c^- lie in the negative region. Now consider the straight lines

$$\sin^2 x \sin^2 y = 1 \tag{6}$$

which intersect the surface (1) at its points of zero curvature. Denote by $g(x, y, z) = 0$ the equation of a surface possessing infinitely many branches and the following properties: (i) each branch is a cylinder with one of the lines of (6) as an axis, (ii) the curvature of such a cylinder is negative in the positive region of x, y, z-space, (iii) the surface admits the group (4) and

intersects the surface $f(x, y, z) = 0$ in small closed curves about the points of intersection of (6) with (1) which do not intersect any of the ovals belonging to the sets $a^+, b^+, c^+, a^-, b^-, c^-$. Let us suppose that $g(x, y, z) > 0$ if the point (x, y, z) lies without the cylinders comprising the surface $g(x, y, z) = 0$ and let us form the sheaf of surfaces

$$fg = \epsilon > 0, \tag{7}$$

retaining only the branch which lies in the region $f > 0$, $g > 0$. From results of Hadamard,* ϵ may be chosen small enough so that the corresponding surface (7) is of negative curvature everywhere. Finally the surface (7) obviously admits the group (4) and may even, by proper choice of the surface $g(x, y, z) = 0$, be made analytic.

If one again designates the portion of the surface (7) contained in the fundamental parallelopiped (5) as the fundamental domain the entire surface is seen as infinitely many copies of the fundamental domain distributed throughout the entire x, y, z-space by the transformations of the group (4). When the boundaries of this fundamental domain are united in the same manner as were the boundaries of the fundamental domain of the surface (1) an abstract, closed surface of genus $p = 2$ and possessing negative curvature everywhere is obtained.

The procedure for constructing an abstract, closed surface of any genus $p > 2$ possessing negative curvature everywhere is now obvious. It will be sufficient to sketch the method for $p = 3$. We begin with a straight line parallel to the z-axis, lying in the plane $y = x$, which intersects the fundamental domain of the surface (7) in two real, distinct points. In place of the straight lines (6) we now employ the straight lines obtained from the above straight line by means of the transformations of the group (4) as axes of the cylinders making up a new surface $h(x, y, z) = 0$ possessing the same properties with respect to the surface (7) as the surface $g(x, y, z) = 0$ possesses with respect to the surface (1). One then forms the surface

$$h(fg - \epsilon) = \delta > 0, \qquad \delta \text{ sufficiently small}$$

and thereby obtains a new surface S permitting the group (4). When the boundaries of the fundamental domain of this surface are joined in the above prescribed manner an abstract, closed surface of genus $p = 3$ and possessing negative curvature everywhere is obtained.

The group (4) and the subgroup formed by the translations IV and V

* J. Hadamard, "Les surfaces à courbures opposées et leur lignes géodesiques," *Journal de Mathématiques*, ser. 5, Vol. 4 (1898), pp. 41-42.

already mentioned are not the only groups taking the surface (1) into itself. There are many others. As an example the transformations of the group

$$(8)\qquad \begin{array}{lll} \bar{x} = x + (2p-1)\pi, & \bar{y} = y + (2q-1)\pi, & \bar{z} = z + (2r-1)\pi, \\ \bar{x} = x + 2k\pi, & \bar{y} = y + 2l\pi, & \bar{z} = z + 2m\pi, \\ & (p, q, r, k, l, m = 0, \pm 1, \cdots) & \end{array}$$

(wherein the members of the upper row are taken together to form a single transformation of the group) take the surface (1) into itself. A fundamental domain for the transformations of this group is the piece of surface contained in the parallelopiped, the equation of whose bounding planes are

$$(9)\qquad y = \pm x \pm \pi, \quad z = \pi \pm \pi/2.$$

It is readily verified that the edges of the parallelopiped (9) which lie in the planes $z = \pi \pm \pi/2$ also lie on the surface (1). Moreover the above edges of this parallelopiped form the boundaries of the fundamental domain. Since the surface (1) permits the group (8) the boundaries of the fundamental domain may be united as follows:

$$\begin{array}{lll} y = x - \pi,\ z = \pi - \pi/2 & : & y = x + \pi,\ z = \pi + \pi/2, \\ y = x - \pi,\ z = \pi + \pi/2 & : & y = x + \pi,\ z = \pi - \pi/2, \\ y = -x - \pi,\ z = \pi - \pi/2 & : & y = -x + \pi,\ z = \pi + \pi/2, \\ y = -x - \pi,\ z = \pi + \pi/2 & : & y = -x + \pi,\ z = \pi - \pi/2, \end{array}$$

the colon denoting a union of two boundaries. One thereby obtains an abstract, closed surface possessing negative curvature except for four points of zero curvature. This abstract, closed surface is non-orientable. In order to see this consider any point (x, y, z) of the fundamental domain and the point $(x + \pi,\ y + \pi,\ z + \pi)$ lying on the surface (1). Join these two points by a simple arc lying on the surface (1). This simple arc appears on the abstract, closed surface as a curve intersecting itself at the point (x, y, z). At each point of the simple arc, considered now as a curve on the surface (1), we erect a normal to the surface (1) directed into the positive region of x, y, z-space. On the abstract, closed surface this corresponds to a continuous displacement of a directed normal from the point (x, y, z) along a curve returning to the point (x, y, z) and it is readily seen that on the return to the point (x, y, z) the sense of the normal is opposite to the sense of the normal with which the displacement was begun.

HARVARD UNIVERSITY.

ON EXISTENCE THEOREMS CONCERNING THE ANALYTICAL TRANSFORMATIONS OF SPACES OF INFINITELY MANY DIMENSIONS INTO THEMSELVES.

By Monroe H. Martin.†

By a power series in infinitely many variables is understood formally an expression of the form

$$\psi \equiv \psi(z_1, z_2, \cdots) \equiv a + \sum_{\nu_1} a_{\nu_1} z_{\nu_1} + \sum_{\nu_1, \nu_2} a_{\nu_1\nu_2} z_{\nu_1} z_{\nu_2} + \cdots,$$

and, by its best majorant, the series ‡

$$\bar{\psi} \equiv \bar{\psi}(z_1, z_2, \cdots) \equiv |a| + \sum_{\nu_1} |a_{\nu_1}| z_{\nu_1} + \sum_{\nu_1, \nu_2} |a_{\nu_1\nu_2}| z_{\nu_1} z_{\nu_2} + \cdots.$$

The transformation of the space of infinitely many dimensions first treated is assumed to be given by

$$x_i = y_i - P_i(y_1, y_2, \cdots), \tag{1}$$

where, by $P_i(y_1, y_2, \cdots)$, we understand

$$P_i(y_1, y_2, \cdots) \equiv \sum_{\nu_1, \nu_2} a^{(i)}_{\nu_1\nu_2} y_{\nu_1} y_{\nu_2} + \sum_{\nu_1, \nu_2, \nu_3} a^{(i)}_{\nu_1\nu_2\nu_3} y_{\nu_1} y_{\nu_2} y_{\nu_3} + \cdots, \tag{1'}$$

the subscripts taking, as is to be understood for all subscripts used in the paper, the values $1, 2, \cdots$. The $P_i(y_1, y_2, \cdots)$ are accordingly power series in the y's with constant coefficients containing no constant and no linear term and we do not require either the variables or the coefficients to be real. In so far as these power series are convergent § the infinite system (1) is said to define an analytical transformation of the y-space into the x-space.

Concerning the infinite system (1) we show the following theorem:

If there exist two positive numbers B and M so that

$$\bar{P}_i(B, B, \cdots) \leqq M \tag{2}$$

then it is possible to find two positive numbers β ($\leqq B$) and γ for which the system (1) *defines in the domain*

$$|x_1| \leqq \gamma, \qquad |x_2| \leqq \gamma, \cdots, \tag{3}$$

† National Research Fellow.

‡ For example $\bar{\sin}\, z = \sinh z$.

§ For the definition of the convergence of a power series in infinitely many variables cf., for example, A. Wintner, "Upon a theory of infinite systems of non-linear implicit and differential equations," *American Journal of Mathematics*, vol. 53 (1931), p. 242.

a uniquely determined inverse transformation

$$y_i = y_i(x_1, x_2, \cdots), \tag{4}$$

which is a power series in the arguments $x_1, x_2, \cdots$ *and is such that if the point in the x-space lies in the complex cube of infinitely many dimensions* (3) *the corresponding point* (4) *in the y-space lies in the complex cube of infinitely many dimensions*

$$| y_1 | \leqq \beta, \quad | y_2 | \leqq \beta, \cdots. \tag{5}$$

A theorem analogous to this theorem follows, as Hilbert † remarks, from some results of Koch. In Hilbert's theorem the power series (1′) are not subjected to the inequalities (2) but there is assumed to exist a sequence of positive numbers $M_1, M_2, \cdots$ for which $\Sigma M_i < +\infty$ and such that

$$| a^{(i)}_{\nu_1\nu_2\ldots} | < M_i. \tag{6}$$

However, our theorem applies to many systems which do not come within the scope of Hilbert's theorem. Examples of such systems are afforded by systems in which the P_i have the same order of magnitude. A typical example is

$$P_i(y_1, y_2, \cdots) \equiv y^2{}_{i+1}.$$

Since the time of Koch there have arisen ‡ very general existence theorems on infinite systems of the form

$$y_i = \lambda f_i \tag{7}$$

where f_i are power series in the y_i, in λ and in a finite or infinite number of parameters x_i. The problem is here to express the variables y_i as functions of λ and the parameters x_i entering in f_i.

In this paper we show how the problem of solving system (1) may be reduced by a simple device to the problem of solving the infinite system (7), the existence theorem for the infinite system (7) yielding our theorem for the infinite system (1). It is then shown that the restriction for the transformation determined by the linear terms to be the identity transformation is not an essential one. For example, it is shown that the infinite system given by

$$x_i = \sum_j c_{ij} y_j - P_i(y_1, y_2, \cdots), \tag{8}$$

in which the infinite matrix is normal in the sense of Koch, i. e.

---

† D. Hilbert, "Wesen und Ziele eine Analysis der unendlichvielen Veränderlichen," *Rendiconti del Circolo Matematico di Palermo*, vol. 27 (1909), p. 73.

‡ A. Wintner, "Differentialgleichungen der Himmelsmechanik," *Mathematische Annalen*, vol. 96 (1927), pp. 291-294.

$$\sum_{i,j} | c_{ij} - \delta_{ij} | < + \infty \tag{9}$$

and the $P_i(y_1, y_2, \cdots)$ are power series in the y_i, as defined in (1′), permits the same existence theorem on the inverse transformation as the system (1). At the end of the paper we indicate how a theorem analogous to the one above may be obtained for infinite systems of the type

$$x_i = y_i - P_i(y_1, y_2, \cdots ;\ x_1, x_2, \cdots),$$

in which the P_i are power series in the y's and the x's for which

$$P_i(0, 0, \cdots ;\ x_1, x_2, \cdots) \equiv 0, \qquad (\partial P_i/\partial y_j)(0, 0, \cdots ;\ x_1, x_2, \cdots) \equiv 0.$$

Corresponding to more general infinite systems there exist more general existence theorems which may be employed to yield existence theorems for more general transformations of the space of infinitely many dimensions into itself. For the sake of clearness, however, we restrict ourselves to the relatively simple system (1).

The existence theorem for an infinite system (7) is as follows: If there exist four positive numbers A, B, C, and D for which the following inequalities are fulfilled:

$$\bar{f}_i(A;\ B, B, \cdots ;\ C, C, \cdots) \leqq D \tag{10}$$

f_i being a power series $f_i(\lambda;\ y_1, y_2, \cdots ;\ x_1, x_2, \cdots)$ in the variables $\lambda;\ y_1, y_2, \cdots$ and the parameters $x_1, x_2, \cdots$, then the infinite system

$$y_i = \lambda f_i$$

possesses in the domain †

$$| \lambda | \leqq \min (A, B/D);\quad | x_1 | \leqq C,\ | x_2 | \leqq C, \cdots, \tag{11}$$

one and only one power series solution

$$y_i(\lambda;\ x_1, x_2, \cdots), \tag{12}$$

and, in the above domain (11) this power series solution satisfies the inequalities

$$| y_i(\lambda;\ x_1, x_2, \cdots) | \leqq B. \tag{13}$$

The device mentioned above for treating the infinite system (1) on the basis of results known for the infinite system (7) consists in introducing an auxiliary parameter λ in the infinite system (1). We obtain thereby a sheaf of infinite systems of the type (1) and we consider, in particular, the sheaf

† We understand by min (a, b) the least of the two numbers a, b for $a \neq b$ and for $a = b$ their common value.

(14) $$y_i = \lambda \Phi_i$$

where we have written

(15) $$\Phi_i \equiv \Phi_i(y_1, y_2, \cdots; x_i) \equiv x_i + P_i(y_1, y_2, \cdots).$$

If we consider the sheaf of infinite systems (14) as an infinite system of the type (7) the functions Φ_i play in the infinite system (14) the same rôle as the functions f_i in the infinite system (7) and corresponding to the inequalities (10) there exist, from (2), three positive numbers B, C, D for which

(16) $$\tilde{\Phi}_i(B, B, \cdots; C) \equiv C + \tilde{P}_i(B, B, \cdots) \leqq D.$$

Since the power series Φ_i do not contain the variable λ the above existence theorem simplifies somewhat and states that in the domain

(17) $$|\lambda| \leqq B/D; \quad |x_1| \leqq C, \quad |x_2| \leqq C, \cdots$$

the infinite system (14) possesses one and only one power series solution $y_i(\lambda; x_1, x_2, \cdots)$ and in the above domain this power series solution satisfies the inequalities

(18) $$|y_i(\lambda; x_1, x_2, \cdots)| \leqq B.$$

It is clear from (16) and the definition of the symbol $\sim$ that if $0 < \beta \leqq B$ and $0 < \gamma \leqq C$ then there exists a $\delta > 0$ for which

(16′) $$\tilde{\Phi}_i(\beta, \beta, \cdots; \gamma) \leqq \delta \leqq D.$$

Accordingly the infinite system (14) possesses in the domain

(17′) $$|\lambda| \leqq \beta/\delta; \quad |x_1| \leqq \gamma, \quad |x_2| \leqq \gamma, \cdots,$$

one and only one power series solution and, in the domain (17′), this power series solution satisfies the inequalities

(18′) $$|y_i(\lambda; x_1, x_2, \cdots)| \leqq \beta.$$

Returning now to the conception of the infinite system (14) as a sheaf of infinite systems of type (1) with parameter λ and including the infinite system (1) as the member of the sheaf corresponding to the parameter value $\lambda = 1$ we shall show that the positive numbers β and γ can be so chosen that the domain

$$|\lambda| \leqq \beta/\delta$$

contains the parameter value $\lambda = 1$, i. e.

(19) $$\beta/\delta > 1.$$

It will thereby be shown that the infinite system (1) possesses a uniquely determined power series inverse transformation as stated in our theorem. From (2) and (1′) we have for any $\beta \leqq B$ the inequality

$$\bar{P}_i(\beta, \beta, \cdots) \leqq \beta^2 p \quad \text{where} \quad p = M/B^2. \tag{20}$$

It follows from (15) and (20) that

$$\bar{\Phi}_i(\beta, \beta, \cdots; \gamma) \leqq \gamma + \beta^2 p,$$

so that we may put in (16′)

$$\delta = \gamma + \beta^2 p.$$

Therefore if we choose β and γ so that

$$\gamma = \beta^2, \qquad \beta < 1/(1+p), \qquad \beta \leqq c^{1/2},$$

we have

$$\beta/\delta = \beta/(\gamma + \beta^2 p) = 1/(1+p)\beta > 1,$$

that is, the inequality (19).

We now take up the case of the apparently more general transformation (8), (9) and shall show how it may, by introduction of new coördinates, be reduced to the case previously treated. We preface this very simple proof by recalling some facts † in connection with matrices which are normal in the sense of Koch. From (9) there obviously follows the existence of a constant K for which

$$\sum_j |c_{ij}| \leqq K. \tag{21}$$

The determinant det $\| c_{ij} \|$ is convergent. If det $\| c_{ij} \|$ is not zero the matrix $\| c_{ij} \|$ has a uniquely determined reciprocal matrix $\| C_{ij} \|$ and there exists a positive number L so that

$$\sum_j |C_{ij}| \leqq L. \tag{22}$$

Accordingly the system of linear equations

$$\sum_j c_{ij} y_j = \eta_i \tag{23}$$

in which the η_i form a bounded sequence possesses one and only one bounded solution, namely

$$y_i = \sum_j C_{ij} \eta_j. \tag{23′}$$

As a matter of fact if $|\eta_i| < \eta$ then $|y_i| < \eta L$. If we introduce new

† Cf., for example, F. Riesz, *Les Systèmes d'équations linéaires à une infinité d'inconnues*, (1913), pp. 24-33.

variables η_i in place of y_i in (8) by means of (23), (23′) the infinite system (8) takes the form

$$x_i = \eta_i - P_i(\sum_j C_{1j}\eta_j, \sum_j C_{2j}\eta_j, \cdots). \tag{24}$$

Corresponding to the domain

$$|y_1| \leqq B^*, \ |y_2| \leqq B^*, \cdots,$$

in the y-coördinate system we obtain the domain

$$|\eta_1| \leqq KB^*, \ |\eta_2| \leqq KB^*, \cdots,$$

in the η-coördinate system and there is a one-to-one correspondence between these two domains. We now write

$$Q_i(\eta_1, \eta_2, \cdots) \equiv P_i(\sum_j C_{1j}\eta_j, \sum_j C_{2j}\eta_j, \cdots)$$

so that (24) takes the form

$$x_i = \eta_i - Q_i(\eta_1, \eta_2, \cdots).$$

In order to complete the proof we need only show the existence of two positive numbers $\bar{B}$ and $\bar{M}$ for which

$$\bar{Q}_i(\bar{B}, \bar{B}, \cdots) \leqq \bar{M}.$$

That this is possible is trivial for we have

$$\bar{Q}_i(|\eta_1|, \ |\eta_2|, \cdots) \leqq \bar{P}_i(\sum_j |C_{1j}| \ |\eta_j|, \ \sum_j |C_{2j}| \ |\eta_j|, \cdots)$$
$$\leqq \bar{P}_i(KLB^*, KLB^*, \cdots)$$

and we may choose $KLB^* \leqq B$.

A word may be desirable here to the effect that since we can find $\bar{\gamma}$, $\bar{\beta} \leqq \bar{B}$, so that $\eta_i \leqq \bar{\beta}$, $\bar{\beta}$ being as small as we please, we can then find γ^*, $\beta^* \leqq B^*$, so that $y_i \leqq \beta^*$.

Our theorem, together with the extension given above, is valid, with a slight change in wording, for the more general infinite system pointed out on page 663. In place of the inequalities (2) the power series P_i are assumed to fulfill the inequalities

$$\bar{P}_i(B, B, \cdots; C, C, \cdots) \leqq M,$$

where B, C and M are given positive numbers; the statement of the remainder of the theorem being unchanged. By hypothesis the power series P_i again contain no terms in which the variables y_i are absent or occur linearly. Consequently the treatment of this more general case proceeds exactly as in the more special case discussed in detail above.

HARVARD UNIVERSITY.

RECURRENCES FOR CERTAIN FUNCTIONS OF PARTITIONS.

By E. T. Bell.

1. *The functions* θ, γ, λ. Let $\theta_1(n) = \Sigma(-1)^t t$, summed over all divisors t of n whose conjugates are odd; $\gamma_1(n) = \Sigma(-1)^d d$, summed over all divisors d of n, and $\lambda_1(n) = \Sigma d$, summed over all divisors d of n. Then, as very special cases of much more general recurrences for functions of divisors,* we have the following.

$$\theta_1(n) + 2\sum_{s=1}\theta_1(n - s^2) = 0 \text{ unless } n = p^2 \quad (p > 0),$$

when the value of the sum is $-p^2$;

$$\sum_{s=0}(-1)^s(2s+1)\lambda_1(n - \tfrac{1}{2}s(s+1)) = 0 \text{ unless } n = \tfrac{1}{2}p(p+1)\,(p > 0),$$

when the value of the sum is $(-1)^{p-1}p(p+1)(2p+1)/6$;

$$\sum_{s=0}\gamma_1(n - \tfrac{1}{2}s(s+1)) = 0 \text{ unless } n = \tfrac{1}{2}p(p+1) \quad (p > 0),$$

when the value of the sum is $-n$. The sums continue so long as the arguments are > 0.

In addition to the generalization mentioned in the footnote (*), above, these recurrences have a curious extension in another direction, namely to the theory of partitions. The above recurrences are the special cases for $r = 1$ of the functions θ_r, γ_r, λ_r of partitions, which are defined as follows.

In MacMahon's suggestive notation for partitions, the partition of n into $a_1 + \cdots + a_r$ parts precisely a_i of which are each equal to n_i $(i = 1, \cdots r)$, is written symbolically $n_1^{a_1} \cdots n_r^{a_r}$, where, without loss of generality, the distinct parts $n_1, \cdots, n_r$ may be considered to be such that $n_1 < \cdots < n_r$. In non-symbolic notation, $n = a_1n_1 + \cdots + a_rn_r$, $n_1 < \cdots < n_r$. To bring out the analogy with divisors we shall use MacMahon's symbolic notation for partitions. In what immediately follows, n_i, n'_i, m_i are integers > 0, the m_i are odd $(i = 1, \cdots, r)$. Two kinds of partitions of n are considered, where r is a constant integer $\geqq 1$,

(1) $$n = m_1^{n_1} \cdots m_r^{n_r}, \qquad m_1 < \cdots < m_r;$$

(2) $$n = n_1'^{n_1} \cdots n_r'^{n_r}, \qquad n'_1 < \cdots < n'_r,$$

* Of the type considered in my paper, *Quarterly Journal*, vol. 49 (1923), pp. 186-192.

and four functions θ_r, ψ_r, γ_r, λ_r are defined for such partitions;

$$\theta_r(n) \equiv \Sigma\,(-1)^{n_1+\cdots+n_r}\, n_1\cdots n_r, \qquad \psi_r(n) \equiv \Sigma\, n_1\cdots n_r,$$

the sum referring to all partitions of type (1) for n fixed;

$$\gamma_r(n) \equiv \Sigma\,(-1)^{n_1+\cdots+n_r}\, n_1\cdots n_r, \qquad \lambda_r(n) \equiv \Sigma\, n_1\cdots n_r,$$

the sum referring to all partitions of type (2) for n fixed. Hence, for $r=1$, θ_1, γ_1, λ_1 are the functions of divisors with which we began.

It will be seen presently that $\psi_r(n) = (-1)^n\,\theta_r(n)$, so that we need discuss only θ_r, γ_r, λ_r. For $|\,q\,| < 1$ the following series are absolutely convergent,

$$\Theta_r(q) \equiv \sum_{n=1}^{\infty} \theta_r(n)q^n, \quad \Gamma_r(q) \equiv \sum_{n=1}^{\infty} \gamma_r(n)q^{2n}, \quad \Lambda_r(n) \equiv \sum_{n=1}^{\infty} \lambda_r(n)q^{2n}.$$

In a different notation, MacMahon * investigated these functions, finding recurrences for them, and thence also recurrences for θ_r, γ_r, λ_r. However, the unsymmetrical development of his algebra, and the fact that he failed to make as full a use of the elementary properties of elliptic theta functions as is at once suggested by the infinite product expansions of these functions, prevented him from finding the simplest recurrences of all, and his results, ill-adapted to computation if $r > 1$, are needlessly complicated. The new recurrences may be stated as follows, bringing out the complete analogy between the cases $r = 1$ and $r > 1$.

(3) $$\theta_r(n) + 2\sum_{s=1} \theta_r(n - s^2) = 0 \text{ if } n < (r+p-1)^2,\ p > 0,$$

or if $n \neq (r+p-1)^2$, while if $n = (r+p-1)^2$ the value of the sum is

$$\frac{2(-1)^r\,(r+p-1)}{2r+p-1}\ \binom{2r+p-1}{p-1}.$$

(4) $$\sum_{s=0} (-1)^s\,(2s+1)\lambda_r(n - \tfrac{1}{2}s(s+1)) = 0 \text{ if } n < r(r+1)/2,$$

or if $n \neq (r+p)(r+p-1)/2$, $p > 0$, while if $n = (r+p)(r+p-1)/2$ the value of the sum is

$$\frac{(-1)^{p-1}}{2^{2r}} \cdot \frac{(4r+2p-2)\,!}{(2r+1)\,!\,(2p-1)\,!}.$$

(5) $$\sum_{s=0} \gamma_r(n - \tfrac{1}{2}s(s+1)) = 0 \text{ if } n < r(r+1)/2,$$

or if $n \neq (r+p)(r+p-1)/2$, $p > 0$, while if $n = (r+p)(r+p-1)/2$ the value of the sum is

* P. A. MacMahon, *Proceedings of the London Mathematical Society*, ser. 2, vol. 19 (1920-21), pp. 75-113.

$$(-1)^r \binom{2r+p-1}{2r}.$$

All of these can be generalized to contain arbitrary arithmetical functions, but for the present we shall prove only (3)-(5).

2. *Proofs of* (3), (4), (5). It will be sufficient to give the formulas from which (3)-(5) follow by simple algebra, with an outline of the main steps.*

Sums and products refer to all $n = 1, 2, 3, \cdots$, or to all $m = 1, 3, 5, \cdots$. We have

$$
\begin{aligned}
&q_0 = \Pi(1-q^{2n}), \qquad q_1 = \Pi(1+q^{2n}),\\
&q_2 = \Pi(1+q^m), \qquad q_3 = \Pi(1-q^m),\\
&\vartheta_0 = q_0 q_3^2, \qquad \vartheta_3 = q_0 q_2^2, \qquad \vartheta'_1 = 2q^{1/4} q_0^3, \qquad \vartheta_2 = 2q^{1/4} q_0 q_1^2;\\
&\vartheta_0(x) = q_0\, \Pi(1 - 2q^m \cos 2x + q^{2m}), \qquad \vartheta_0 = 1 + 2\,\Sigma\,(-1)^n q^{n^2},\\
&\vartheta_3(x) = q_0\, \Pi(1 + 2q^m \cos 2x + q^{2m}), \qquad \vartheta_3 = 1 + 2\,\Sigma\, q^{n^2},\\
&\csc x \vartheta_1(x) = 2q^{1/4} q_0\, \Pi(1 - 2q^{2n} \cos 2x + q^{4n}), \qquad \vartheta'_1 = 2\,\Sigma(-1 \mid m)\, m q^{m^2/4},\\
&\sec x \vartheta_2(x) = 2q^{1/4} q_0\, \Pi(1 + 2q^{2n} \cos 2x + q^{4n}), \qquad \vartheta_2 = 2\,\Sigma\, q^{m^2/4},
\end{aligned}
$$

where $(-1 \mid m) \equiv (-1)^{(m-1)/2}$.

Let $C \equiv c_1, c_2, \cdots, 0 < c_1 < c_2 < \cdots$, be an infinite ascending sequence of integers > 0, and let $f(x)$ be single-valued and finite for integer values > 0 of x. Then, formally,

$$\prod_{r=1}^{\infty} (1 + f(c_r)y) = 1 + \sum_{r=1}^{\infty} y^r (\Sigma f(c_p) \cdots f(c_q)),$$

where the Σ refers to *all* sets of r elements $c_p, \cdots, c_q$ of C which are such that $c_p < \cdots < c_q$. Take $f(x) \equiv kq^x/(1-aq^x)^2$, and expand formally. Then, for this $f(x)$, we have

$$\prod_{r=1}^{\infty} (1 + f(c_r)y) = 1 + \sum_{r=1}^{\infty} k^r y^r \left[\sum_{n=1}^{\infty} q^n S(n)\right],$$

$$S(n) \equiv \Sigma\, n_1 \cdots n_r\, a^{n_1 + \cdots + n_r - r},$$

the sum referring to all solutions of

$$n = n_1 c_{p_1} + \cdots + n_r c_{p_r}, \quad 0 < c_{p_1} < \cdots < c_{p_r}, \quad 0 < n_i \quad (i = 1, \cdots, r).$$

Apply the last to the product expansions of the thetas, using the identities

* The formulas required from the theta functions and constants q_α are summarized in Tannery and Molk's *Éléments*, vol. 2, pp. 252, 257; the reduction formulas from trigonometry are given in Hobson's treatise, chap. 7.

$$1 \pm 2q^h \cos 2x + q^{2h}$$

$$= (1 \mp q^h)^2 \left[1 \pm \frac{4q^h \cos^2 x}{(1 \mp q^h)^2}\right] = (1 \pm q^h)^2 \left[1 \mp \frac{4q^h \sin^2 x}{(1 \pm q^h)^2}\right].$$

Write $\Psi_r(q) = \Sigma\, \psi_r(n) q^n$, $z = 2\cos x$, $w = 2\sin x$. Then, the Σ in the following referring to $r = 1$, $r = \infty$, we have

$$\begin{aligned}
\vartheta_0(x) &= \vartheta_3[1 + \Sigma\, \Theta_r(q) z^{2r}] &&= \vartheta_0[1 + \Sigma\, \Psi_r(q) w^{2r}],\\
\vartheta_3(x) &= \vartheta_0[1 + \Sigma\, \Psi_r(q) z^{2r}] &&= \vartheta_3[1 + \Sigma\, \Theta_r(q) w^{2r}],\\
\csc x \vartheta_1(x) &= \vartheta_2[1 + \Sigma\, \Gamma_r(q) z^{2r}] &&= \vartheta'_1[1 + \Sigma\, \Lambda_r(q) w^{2r}],\\
\sec x \vartheta_2(x) &= \vartheta'_1[1 + \Sigma\, \Lambda_r(q) z^{2r}] &&= \vartheta_2[1 + \Sigma\, \Gamma_r(q) w^{2r}].
\end{aligned}$$

Since $\vartheta_0(x, -q) = \vartheta_3(x, q)$, etc., we have

$$\Psi_r(q) = \Theta_r(-q), \qquad \Gamma_r(q) = \Gamma_r(-q), \qquad \Lambda_r(q) = \Lambda_r(-q);$$

from the first of which, $\psi_r(n) = (-1)^n \theta_r(n)$; the second and third merely provide slight checks on the algebra.

In the Fourier series for the thetas,

$$\vartheta_3(x) = 1 + 2\,\Sigma\, q^{n^2} \cos 2nx, \qquad \vartheta_1(x) = 2\,\Sigma\, (-1 \mid m) q^{m^2/4} \sin mx,$$

$$\sec x \vartheta_2(x) = 2\,\Sigma\, q^{m^2/4} \cos mx \sec x \quad (n = 1, 2, 3, \cdots;\ m = 1, 3, 5, \cdots),$$

expand the $\cos 2nx$, $\sin mx$, $\cos mx \sec x$ into polynomials in w (w as above), using the following easily obtained reductions of the standard formulas,

$$\cos 2nx = 1 + n \sum_{s=1}^{n} \frac{(-1)^s (n + s - 1)!\, 2^{2s}}{(2s)!\, (n - s)!} w^{2s},$$

$$\sin mx = m \sin x + \sum_{s=2}^{(m+1)/2} \frac{(-1)^{s-1} (m + 2s - 3)!}{(2s - 1)!\, (m - 2s + 2)!\, 2^{2s-1}} w^{2s-1} \ (m > 1),$$

$$\cos mx \sec x = 1 + \sum_{s=1}^{(m-1)/2} \frac{(-1)^s \left(\frac{m + 2s - 1}{2}\right)!\, 2^{2s}}{(2s)! \left(\frac{m - 2s - 1}{2}\right)!} w^{2s} \ (m > 1).$$

Equating coefficients of like powers of q after substituting these into $\vartheta_\alpha(x)$ ($\alpha = 3, 2, 1$) in the identities containing Θ_r, Λ_r, Γ_r, we find (3)-(5) of § 1.

CONSTRUCTION OF TRANSFORMATIONS TO CANONICAL FORMS.

By Walter O. Menge.

1. *Introduction.* The problem of establishing the existence of a transformation to the rational canonical form has received considerable attention from various writers, each of whom has provided a characterization of the form. At least two of these have submitted existence proofs which are satisfactory from a modern standpoint. The first is the classical proof of Frobenius [*] and the second is due to Dickson and his predecessors and completed at an essential point by Bennett.[†] Each of these proofs employs only rational operations in the field of the coefficients of the original form and each is constructible in the sense that all the operations may be carried out in any special case in a finite number of steps. This paper therefore assumes the following:

Theorem 1. *Let A be any n-rowed square matrix with elements in any field F. Then A is similar to a matrix P_r of the rational canonical form, that is, there exists a non-singular matrix B, with elements in F, such that*

$$B \cdot A \cdot B^{-1} = P_r.$$

Although the two proofs mentioned above afford constructions of the matrix B in a finite number of steps, the construction is extraordinarily laborious by either method. The purpose of this paper is to present a solution of the problem of finding a practicable method of determining the matrix B in the above theorem, given a particular A and hence a particular P_r.

In the latter part of the paper there is presented explicitly a transformation from the rational canonical form to the classic canonical form, and another for the reverse process. These forms for the transformation possess the advantage of directness and furnish simple practical schemes under entirely general hypotheses. Numerical examples are appended for the purpose of illustrating the method.

2. *Rational canonical form.* Consider the linear forms

$$(1) \qquad X_i = \sum_{j=1}^{n} a_{ij} x_j \qquad (i = 1, 2, \cdots, n)$$

[*] Muth, *Elementartheiler* 3, and references given there.

[†] L. Dickson, "Modern algebraic theories," Chapter V; A. Bennett, *American Mathematical Monthly*, vol. 38, pp. 377-383.

in the independent variables $x_1, x_2, \cdots, x_n$ with constant coefficients a_{ij} in any field F. The case in which the matrix A of the coefficients is singular is not excluded, so that the transformation (1) is entirely general.

Designate by $G_j(\lambda)$ the highest common factor of all j-th minors (of order $n-j$) of the characteristic determinant

$$(2) \qquad D(\lambda) = | A - \lambda I |.$$

These common factors are chosen so that the coefficient of the highest power of the variable λ is unity. The invariant factors of the matrix A are defined as

$$(3) \qquad \begin{aligned} D_j(\lambda) &= G_{j-1}(\lambda)/G_j(\lambda) \\ &= \lambda^{n_j} - d_{1j} - d_{2j}\lambda - \cdots - d_{n_j j}\lambda^{n_j-1}, \qquad (j = 1, 2, \cdots, n). \end{aligned}$$

Define the integer σ, so that

$$D_j(\lambda) \neq 1 \text{ for } j \leq \sigma; \qquad D_j(\lambda) = 1 \text{ for } j > \sigma.$$

It can be shown * that if transformation (1) be subjected to the introduction of new variables,

$$(4) \qquad \begin{aligned} y_i &= \sum_{j=1}^{n} b_{ij}x_j \qquad (i = 1, 2, \cdots, n), \\ Y_i &= \sum_{j=1}^{n} b_{ij}X_j \qquad (i = 1, 2, \cdots, n), \end{aligned}$$

(the matrix B of the coefficients b_{ij} being non-singular) then the matrix P of the coefficients p_{ij} in the linear form

$$(5) \qquad Y_i = \sum_{j=1}^{n} p_{ij}y_j \qquad (i = 1, 2, \cdots, n)$$

satisfies the equation

$$(6) \qquad P = B\,A\,B^{-1}.$$

In the special case, when the linear form (5) simplifies to

$$(7) \qquad \begin{aligned} Y_{1j} &= y_{2j} \\ Y_{2j} &= y_{3j} \\ &\cdots\cdots \\ Y_{n_j-1,j} &= y_{n_j j} \\ Y_{n_j j} &= \sum_{s=1}^{n_j} d_{sj}y_{sj} \qquad (j = 1, 2, \cdots, \sigma), \end{aligned}$$

the linear forms are said to be in the rational canonical form. This char-

---

* Dickson, *ibid.*

acterization is equivalent to that of Dickson. Each of the groups displayed in (7), corresponding to a particular invariant factor of the matrix A, is called a "chain." The coefficients d_{s_j} in the last equation are precisely the coefficients of the several powers of the variable λ in the definition (3) of the invariant factors.

The matrix P_r of the coefficients of the rational canonical form (7) is

$$(8)\qquad P_r = \left\| \begin{matrix} M_1 & 0 & 0 & \cdots & 0 \\ 0 & M_2 & 0 & \cdots & 0 \\ 0 & 0 & M_3 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & M_\sigma \end{matrix} \right\|$$

where

$$(9)\qquad M_j = \left\| \begin{matrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ d_{1j} & d_{2j} & d_{3j} & d_{4j} & \cdots & d_{n_j j} \end{matrix} \right\| \qquad (j = 1, 2, \cdots, \sigma).$$

Certain necessary conditions on the coefficients of the transformation (4) can be found by multiplying equation (6) on the right by B, yielding

$$(10)\qquad B \cdot A = P_r \cdot B.$$

Upon equating corresponding elements in the two products comprising the members of this equation one obtains the following necessary conditions to be satisfied by the elements of the matrix B:

(for the first row)

$$(11)\qquad \sum_{s=1}^{n} a_{sk} b_{1s} = b_{2k}, \qquad (k = 1, 2, \cdots, n),$$

(for the second row)

$$(12)\qquad \sum_{s=1}^{n} a_{sk} b_{2s} = b_{3k}, \qquad (k = 1, 2, \cdots, n),$$

. . . .

(for the $n_1 - 1$-th row)

$$(13)\qquad \sum_{s=1}^{n} a_{sk} b_{n_1-1,s} = b_{n_1 k}, \qquad (k = 1, 2, \cdots, n),$$

(for the n_1-th row)

$$(14)\qquad \sum_{s=1}^{n} a_{sk} b_{n_1 s} = \sum_{s=1}^{n} d_{s1} b_{sk}, \qquad (k = 1, 2, \cdots, n),$$

together with systems of equations similar in form to (11), (12), $\cdots$, (13), and (14) corresponding to each of the other chains.

Let B_{i1} $(i=1,2,\cdots,n_1)$ denote respectively the one-rowed matrices formed by the first n_1 successive rows of the matrix B. Let B_{i2} $(i=1,2,\cdots,n_2)$ denote respectively the one-rowed matrices formed by the next n_2 successive rows of the matrix B, etc. Thus, in general

$$B_{ij}=(b_{s1},b_{s2},\cdots,b_{sn}) \tag{15}$$

where

$$s=\sum_{h=1}^{j-1}n_h+i.$$

Equations (11), (12), $\cdots$ (13), and (14) can now be written in the simple form

$$\begin{aligned} B_{11}A&=B_{21}\\ B_{21}A&=B_{31}\\ &\cdots\cdots\\ B_{n_1-1,1}A&=B_{n_11}\\ B_{n_11}A&=\sum_{s=1}^{n_1}d_{s1}B_{s1}. \end{aligned} \tag{16}$$

The systems of equations for the succeeding chains corresponding to the second and succeeding invariant factors would appear as follows:

$$\begin{aligned} B_{1j}A&=B_{2j}\\ B_{2j}A&=B_{3j}\\ &\cdots\cdots\\ B_{n_j-1,j}A&=B_{n_jj}\\ B_{n_jj}A&=\sum_{s=1}^{n_j}d_{sj}B_{sj}. \end{aligned} \qquad (j=2,3,\cdots,\sigma) \tag{17}$$

After the substitution of the values given in each equation of (16) and (17) into the succeeding equations in the same chain, one obtains

$$\begin{aligned} B_{1j}A&=B_{2j}\\ B_{1j}A^2&=B_{3j}\\ &\cdots\cdots\\ B_{1j}A^{n_j-1}&=B_{n_jj}\\ B_{1j}\cdot D_j(A)&=0. \end{aligned} \qquad (j=1,2,\cdots,\sigma), \tag{18}$$

where $D_j(A)$ is the matrix obtained by replacing the variable λ by the matrix A in the j-th invariant factor $D_j(\lambda)$.

When the notation B_{1j} is replaced by L_j $(j=1,2,\cdots,\sigma)$ it appears from equations (18) that the "leaders" L_j of the several chains must necessarily satisfy the several invariant factor equations

$$(19) \qquad L_j \cdot D_j(A) = 0, \qquad (j = 1, 2, \cdots, \sigma).$$

Furthermore, by means of equations (18) every row of the matrix B is expressible in terms of the leader of the particular chain to which that row belongs. With the aid of equations (18) the successive rows of the matrix B can now be written sequentially

$$(20) \qquad L_1 I,\ L_1 A,\ L_1 A^2, \cdots, L_1 A^{n_1-1},\ L_2 I,\ L_2 A, \cdots, L_\sigma I, \cdots, L_\sigma A^{n_\sigma-1}.$$

Since equations (11), (12), $\cdots$, (13), and (14), and consequently equations (18) are necessary conditions on the elements of the matrix B it follows that every transformation to the rational canonical form is of the form (20). It remains to be shown how the leaders L_j $(j = 1, 2, \cdots, \sigma)$ can be so chosen as to produce a non-singular matrix B.

Denote the elements of the one-rowed matrix L_j by l_{ij} $(i = 1, 2, \cdots, n)$ so that

$$(21) \qquad L_j = (l_{1j}, l_{2j}, \cdots, l_{nj}) \qquad (j = 1, 2, \cdots, \sigma).$$

The matrix B can now be exhibited in the form:

$$(22) \qquad B = \left\| \begin{array}{cccc} l_{11} & l_{21} & \cdots & l_{n1} \\ \Sigma\, a_{i1}^{(1)}\, l_{i1} & \Sigma\, a_{i2}^{(1)}\, l_{i1} & \cdots & \Sigma\, a_{in}^{(1)}\, l_{i1} \\ \Sigma\, a_{i1}^{(2)}\, l_{i1} & \Sigma\, a_{i2}^{(2)}\, l_{i1} & \cdots & \Sigma\, a_{in}^{(2)}\, l_{i1} \\ \cdot & \cdot & \cdots & \cdot \\ \Sigma\, a_{i1}^{(n_1-1)}\, l_{i1} & \Sigma\, a_{i2}^{(n_1-1)}\, l_{i1} & \cdots & \Sigma\, a_{in}^{(n_1-1)}\, l_{i1} \\ l_{12} & l_{22} & \cdots & l_{n2} \\ \Sigma\, a_{i1}^{(1)}\, l_{i2} & \Sigma\, a_{i2}^{(1)}\, l_{i2} & \cdots & \Sigma\, a_{in}^{(1)}\, l_{i2} \\ \cdot & \cdot & \cdots & \cdot \\ \Sigma\, a_{i1}^{(n_\sigma-1)}\, l_{i\sigma} & \Sigma\, a_{i2}^{(n_\sigma-1)}\, l_{i\sigma} & \cdots & \Sigma\, a_{in}^{(n_\sigma-1)}\, l_{i\sigma} \end{array} \right\|$$

in which each summation extends from $i = 1$ to $i = n$ and $a_{ik}^{(h)}$ is the element in the i-th row and k-th column of the matrix A^h.

Each of the equations

$$(19) \qquad L_j \cdot D_j(A) = 0 \qquad (j = 1, 2, \cdots, \sigma),$$

can be considered as a set of n linear homogeneous equations in the field F between the elements l_{ij} of the leader L_j. Since the rank of $D_1(A)$ is zero,* the elements of the first leader L_1 can be considered as independent. Denote the ranks of the matrices $D_j(A)$ $(j = 2, 3, \cdots, \sigma)$ by R_j $(< n)$. Hence R_j of the elements in the leader L_j can be determined as linear homogeneous functions, with coefficients in the field F, of the remaining $n - R_j$ elements.

* Menge, "On the rank of the product, etc.," *Bulletin, American Mathematical Society*, Feb. 1932, pp. 88-94.

Designate the remaining $n - R_j$ elements by m_{ij} $(i = 1, 2, \cdots, n - R_j)$. Then each element of the leader L_j is expressible as a linear combination of the independent parameters m_{ij}. The determinant of the matrix B yields upon expansion a homogeneous polynomial of the n-th degree in the parameters m_{ij}. Since it has been shown that a non-singular matrix B exists, there will be at least one term in the expansion of the determinant with a non-zero coefficient. Suppose that such a term is

$$(20) \qquad c \cdot m_{i_1j_1}^{e_1} \cdot m_{i_2j_2}^{e_2} \cdots m_{i_kj_k}^{e_k}, \quad \text{where} \quad \sum_{h=1}^{k} e_{i_hj_h} = n,$$

and c is a numerical coefficient not equal to zero. It should be noted that it may not be necessary to carry out the expansion of the determinant of B in order to find such a term. In most numerical examples this term can be found by inspection. Set

$$m_{i_2j_2} = m_{i_3j_3} = \cdots = m_{i_kj_k} = 1$$

and all of the other parameters involved in the expansion of the determinant of B, with the single exception of $m_{i_1j_1}$, equal to zero. After these substitutions the determinant of the matrix B simplifies to a polynomial, $p(m_{i_1j_1})$, in the single parameter $m_{i_1j_1}$. The non-singularity of the transformation B can now be assured by assigning to the parameter $m_{i_1j_1}$ a value distinct from the several roots of the equation

$$p(m_{i_1j_1}) = 0.$$

A numerical example illustrating the method is appended to this paper.

3. *Classic canonical form.* The second problem to be considered is the explicit representation of the transformation from the rational canonical form to the classic canonical form. Let the invariant factors of the matrix A, and hence also the invariant factors of the matrix P_r, be exhibited as

$$(21) \quad D_j(\lambda) = (\lambda - \lambda_1)^{n_j^{(1)}}(\lambda - \lambda_2)^{n_j^{(2)}} \cdots (\lambda - \lambda_\omega)^{n_j^{(\omega)}}, \quad (j = 1, 2, \cdots, \sigma).$$

When a set of linear forms simplifies to the form

$$(22) \qquad \begin{aligned} Y_{1ij} &= \lambda_i y_{1ij} \\ Y_{2ij} &= \lambda_i y_{2ij} + y_{1ij} \\ Y_{3ij} &= \lambda_i y_{3ij} + y_{2ij} \\ &\cdots\cdots\cdots\cdots \\ Y_{n_j^{(i)}ij} &= \lambda_i y_{n_j^{(i)}ij} + y_{n_j^{(i)}-1,ij} \\ (i = 1, 2, \cdots, \omega;\ & j = 1, 2, \cdots, \sigma), \end{aligned}$$

the linear forms are said to be in the classic canonical form. As is seen from this definition, there is a chain similar in structure to that exhibited above

for every distinct root in every invariant factor of the matrix A. Exhibited as a matrix the coefficients of the classic canonical form (22) appear as

$$(23) \qquad P_c = \left\| \begin{matrix} N_{11} & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & N_{21} & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & N_{n\sigma,\sigma} \end{matrix} \right\|,$$

where N_{ij} is an $n_j^{(i)}$-rowed square matrix of the form

$$(24) \qquad N_{ij} = \left\| \begin{matrix} \lambda_i & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 1 & \lambda_i & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & \lambda_i & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \lambda_i \end{matrix} \right\|$$
$$(i = 1, 2, \cdots, \omega;\ j = 1, 2, \cdots, \sigma).$$

The classic canonical form can be considered from one viewpoint as a further reduction of the rational canonical form, in which each invariant factor chain is sub-divided into smaller chains corresponding to the several elementary divisors comprising the invariant factor. A natural extension of the method of the preceding section would involve the consideration of a method of constructing a non-singular matrix T which would transform the rational canonical form P_r into its classic canonical form P_c. More specifically, the problem at hand is the determination of a non-singular matrix T, such that

$$T \cdot P_r \cdot T^{-1} = P_c.$$

After multiplying this equation on the right by T, one has

$$(25) \qquad T \cdot P_r = P_c \cdot T.$$

In order to effect simplicity in notation consideration at the outset will be limited to the part of the matrix T, denoted by T_j, which will convert the rational canonical form (7) corresponding to the j-th invariant factor into the classic canonical form defined by (22). The equation corresponding to (25) but referring only to the j-th invariant factor is

$$(26) \qquad T_j \cdot M_j = Q_j \cdot T_j,$$

where

$$Q_j = \left\| \begin{matrix} N_{1j} & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & N_{2j} & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & N_{n_j j} \end{matrix} \right\|.$$

Define $$\alpha_i = \sum_{k=1}^{i-1} n_j^{(k)} + 1, \qquad (i = 1, 2, \cdots, \omega).$$

By equating corresponding elements in the two products comprising the members of equation (26) one obtains the following necessary conditions to be satisfied by the elements t_{kl} of the matrix T_j. From the k-th row ($k = 1, 2, \cdots, n_j$) one has

$$\left.\begin{aligned} &(27) & d_{1j}t_{\alpha_i n_j} &= \lambda_i t_{\alpha_i 1} \\ &(28) & t_{\alpha_i, l-1} + d_{lj}t_{\alpha_i n_j} &= \lambda_i t_{\alpha_i l} \end{aligned}\right\} \qquad (k = \alpha_i;\; l = 2, 3, \cdots, n_j),$$

and

$$(29) \qquad d_{1j}t_{kn_j} = t_{k-1,1} + \lambda_i t_{k1}$$

$$(30) \qquad t_{k,l-1} + d_{lj}t_{kn_j} = t_{k-1,l} + \lambda_i t_{kl} \qquad (\alpha_i < k < \alpha_{i+1};\; l = 2, 3, \cdots, n_j).$$

For the rows in which $k = \alpha_i$ ($i = 1, 2, \cdots, \omega$) equations (28) can be solved sequentially beginning with the last for the values of $t_{\alpha_i l}$ ($l = 1, 2, \cdots, n_j - 1$) in terms of $t_{\alpha_i n_j}$. In this manner one finds

$$(31) \qquad t_{\alpha_i l} = g_l(\lambda_i)t_{\alpha_i n_j} \qquad (l = 1, 2, \cdots, n_j;\; i = 1, 2, \cdots, \omega)$$

where

$$(32) \qquad g_l(\lambda_i) = \lambda_i^{n_j - l} - \sum_{h=0}^{n_j-l-1} d_{h+l+1}\lambda_i^h.$$

In deriving equations (31) the properties expressed by equations (27) were not employed, but it may be easily verified that these equations are satisfied identically in the elements $t_{\alpha_i n_j}$. Upon replacing in equations (27) the elements $t_{\alpha_i l}$ by their values as given by equations (31) one obtains

$$(33) \qquad t_{\alpha_i n_j}D_j(\lambda_i) = 0.$$

Equations (33) are identities in the elements $t_{\alpha_i n_j}$ by virtue of the hypotheses regarding the roots of the invariant factors as given in (21). Thus equations (31) and (33) and hence also (27) and (28) are satisfied for any values of the elements $t_{\alpha_i n_j}$ ($i = 1, 2, \cdots, \omega$), and these elements may be regarded as parameters in terms of which all the elements in their rows of the matrix T_j can be expressed.

After solving in the same manner equations (30) for each element t_{kl}, one finds

$$(34) \qquad t_{kl} = \sum_{p=0}^{m} \frac{1}{(m-p)!} g_l^{(m-p)}(\lambda_i)t_{\alpha_i+p, n_j},$$

where $\alpha_i < k < \alpha_{i+1}$; $m = k - \alpha_i$; $i = 1, 2, \cdots, \omega$, ard

$$g_l^{(h)}(\lambda_i) = \left.\frac{d^h g_l(\lambda)}{d\lambda^h}\right]_{\lambda=\lambda_i}.$$

Equations (29) were not employed in obtaining (34). However it may be shown that these equations are also satisfied identically in the elements t_{kn_j} by the expressions given in equations (34). Upon replacing the elements t_{kl} in equations (29) by the values given in equations (34) one finds after simplification

$$(35) \quad t_{kn_j} \cdot D_j(\lambda_1) + t_{k-1,n_j} \cdot D_j^{(1)}(\lambda_1) + \cdots + t_{\alpha_i n_j} \cdot D_j^{(m)}(\lambda_i) = 0.$$

Since λ_i is a root of the equation

$$D_j(\lambda) = 0$$

of multiplicity $n_j^{(i)} = \alpha_{i+1} - \alpha_i$, it follows immediately that the coefficients of t_{hn_j} $(h = \alpha_i, \alpha_i + 1, \cdots, k)$ in equations (35) vanish. Hence (34) and (35) and hence (29) and (30) are satisfied for any values of the elements t_{hn_j} $(h = 1, 2, \cdots, n_j)$ and these elements may be considered as parameters in terms of which all other elements in the same row of T_j can be expressed.

Since equations (34) for the particular values $k = \alpha_i$, $(i = 1, 2, \cdots, \omega)$, reproduce equations (32), one can now write the matrix T_j in the form

$$T_j = (t_{kl}) = \left(\sum_{p=0}^{m} \frac{1}{(m-p)!} g_l^{(m-p)}(\lambda_i) t_{\alpha_i+p,n_j} \right)$$

where $\quad m = k - \alpha_i; \quad \alpha_i < k < \alpha_{i+1},$

in which the elements t_{kn_j} are to be considered as parameters and may have any values which insure the non-singularity of the matrix T_j.

Let

$$(36) \qquad \begin{cases} t_{kn_j} = 1, & \text{for} \quad k = \alpha_i;\ i = 1, 2, \cdots, \omega, \\ t_{kn_j} = 0, & \text{for} \quad k \neq \alpha_i;\ i = 1, 2, \cdots, \omega. \end{cases}$$

For this particular choice of the parameters the matrix T_j takes the special form

$$(37) \qquad T_j = (t_{kl}) = \left(\frac{1}{m!} g_l^{(m)}(\lambda_i) \right)$$

where $m = k - \alpha_i$, $\alpha_i < k < \alpha_{i+1}$; $i = 1, 2, \cdots, \omega$.

It remains to be shown that T_j as expressed in form (37) is non-singular. After the functions $g_l^{(m)}(\lambda_i)$ have been replaced by polynomials in λ_i according to the definitions (32) each element of the l-th column $(l = 1, 2, \cdots, n_j)$ of the determinant of (37) is expressed as the sum of $n_j - l + 1$ terms. The determinant of (37) can then be expanded as the sum of $n_j(n_j + 1)/2$ determinants, each element of each determinant involving one term taken from the corresponding element of the determinant of (37). After removing common factors it is apparent that the new determinants found in

this manner are all zero (by virtue of possessing two identical columns) except the single determinant

$$(38) \qquad \left| \frac{(\lambda_i^{n_j-l})^{(m)}}{m!} \right| .$$

The determinant (38) has been shown * to be non-singular and hence the matrix T_j as expressed explicitly by (37) is non-singular.

It is obvious that the transformation,

$$(39) \qquad T = \begin{Vmatrix} T_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & T_2 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & T_\sigma \end{Vmatrix},$$

in which each submatrix T_j is of the form (37), is non-singular and satisfies the matrix equation

$$T \cdot P_r \cdot T^{-1} = P_c.$$

It follows immediately that the matrix $C = T \cdot B$ is non-singular and satisfies the matrix equation

$$C \cdot A \cdot C^{-1} = P_c,$$

and that the matrix $S = T^{-1}$ satisfies the matrix equation

$$S \cdot P_c \cdot S^{-1} = P_r.$$

A numerical example of the method of this section is appended.

4. *Numerical illustration.* In order to illustrate the method of the previous section a numerical example of the transformations previously derived in general form, will not be presented. Let the given linear form † be

$$(A) \qquad \begin{aligned} X_1 &= -2x_1 - x_2 - x_3 + 3x_4 + 2x_5 \\ X_2 &= -4x_1 + x_2 - x_3 + 3x_4 + 2x_5 \\ X_3 &= x_1 + x_2 \qquad - 3x_4 - 2x_5 \\ X_4 &= -4x_1 - 2x_2 - x_3 + 5x_4 + x_5 \\ X_5 &= 4x_1 + x_2 + x_3 - 3x_4. \end{aligned}$$

After evaluating the characteristic determinant of the coefficients in the linear forms above, one finds

$$| A - \lambda I | = D(\lambda) = (\lambda - 2)^3 (\lambda + 1)^2.$$

* Nyswander, "A direct solution of systems of linear differential equations, etc," *American Journal of Mathematics*, vol. 47 (1925), pp. 272-3.

† Burnside, *Proceedings of the London Mathematical Society*, vol. 30 (1899), pp. 191-2.

The ranks of the determinants $|A-2I|$ and $|A+I|$ are 3 and 4, respectively. It follows immediately * that the invariant factors of the matrix A must be, respectively,

$$D_1(\lambda)=(\lambda-2)^2(\lambda+1)^2=\lambda^4+4+4\lambda-3\lambda^2-2\lambda^3, \qquad D_2(\lambda)=\lambda-2.$$

The matrix equation,

$$L_2D_2(A)=0,$$

to be satisfied by the leader of the second chain, can now be written in the form (dropping the second subscript)

$$\begin{aligned}
-4l_1-4l_2+l_3-4l_4+4l_5&=0\\
-l_1-l_2+l_3-2l_4+l_5&=0\\
-l_1-l_2-2l_3-l_4+l_5&=0\\
+3l_1+3l_2-3l_3+3l_4-3l_5&=0\\
+2l_1+2l_2-2l_3+l_4-2l_5&=0.
\end{aligned}$$

These simultaneous equations have the solution

$$l_1=p_{12};\ l_2=p_{22};\ l_3=0;\ l_4=0;\ l_5=p_{12}+p_{22},$$

where p_{12} and p_{22} are arbitrary parameters.

The transpose (written for convenience) of the matrix B can now be exhibited in the form

$$\left\|\begin{matrix}
p_{11}, & -2p_{11}-4p_{21}+p_{31}-4p_{41}+4p_{51}, & 3p_{11}-p_{21}-2p_{31}-p_{41}+p_{51}, & -4p_{11}-12p_{21}+3p_{31}-12p_{41}+12p_{51}, & p_{12}\\
p_{21}, & -p_{11}+p_{21}+p_{31}-2p_{41}+p_{51}, & -4p_{11}+4p_{31}-8p_{41}+4p_{51}, & -12p_{11}-4p_{21}+12p_{31}-24p_{41}+12p_{51}, & p_{22}\\
p_{31}, & -p_{11}-p_{21}-p_{41}+p_{51}, & 2p_{11}+2p_{21}-p_{31}+2p_{41}-2p_{51}, & -3p_{11}-3p_{21}+2p_{31}-3p_{41}+3p_{51}, & 0\\
p_{41}, & 3p_{11}+3p_{21}-3p_{31}+5p_{41}-3p_{51}, & 3p_{11}+3p_{21}-3p_{31}+7p_{41}-3p_{51}, & 9p_{11}+9p_{21}-9p_{31}+17p_{41}-9p_{51}, & 0\\
p_{51}, & 2p_{11}+2p_{21}-2p_{31}+p_{41}, & -p_{11}-p_{21}+p_{31}-5p_{41}+5p_{51}, & -3p_{11}-3p_{21}+3p_{31}-15p_{41}+11p_{51}, & p_{12}+p_{22}
\end{matrix}\right\|$$

in which all of the parameters p_{ij} are arbitrary, subject only to the condition that the matrix is non-singular. The coefficient of the term $p_{11}{}^4\cdot p_{12}$ in the determinant of the above matrix is 81 and hence the matrix B will be non-singular if $p_{11}=p_{12}=1$, and $p_{21}=p_{31}=p_{41}=p_{51}=p_{22}=0$. After these substitutions the matrix B becomes

* Menge, *ibid.*

$$(B) = \left\| \begin{array}{rrrrr} 1 & 0 & 0 & 0 & 0 \\ -2 & -1 & -1 & 3 & 2 \\ 3 & -4 & 2 & 3 & -1 \\ -4 & -12 & -3 & 9 & -3 \\ 1 & 0 & 0 & 0 & 1 \end{array} \right\|, \text{ and } B \cdot A \cdot B^{-1} = \left\| \begin{array}{rrrrr} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -4 & -4 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array} \right\|,$$

the last matrix being the required rational canonical form of A.

Proceeding to the problem of determining the transformation T from the rational canonical form to the classic canonical form, one finds

$$\begin{aligned} D_1(\lambda) &= \lambda^4 + 4 + 4\lambda - 3\lambda^2 - 2\lambda^3 \\ g_1(\lambda) &= \lambda^3 + 4 - 3\lambda - 2\lambda^2 \\ g_2(\lambda) &= \lambda^2 - 3 - 2\lambda \\ g_3(\lambda) &= \lambda - 2 \\ g_4(\lambda) &= 1. \end{aligned}$$

After evaluating these polynomials and their first derivatives for $\lambda = 2$ and -1, respectively, it is possible to write down immediately the transformation T,

$$(T) = \left\| \begin{array}{rrrr|r} -2 & -3 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ \hline 4 & 0 & -3 & 1 & 0 \\ 4 & -4 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array} \right\|, \text{ and } T \cdot P_r \cdot T^{-1} = \left\| \begin{array}{rrrrr} 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{array} \right\| = (P_c)$$

UNIVERSITY OF MICHIGAN.

* Added in reading proof: There has been recently published another proof of the theorem mentioned in the introduction, see Ingraham, *Bulletin of the American Mathematical Society*, vol. 39 (1933), pp. 379-382.

A CLASS OF REPRESENTATIONS OF MANIFOLDS. PART I.*

By Charles B. Morrey, Jr.†

Since Schwarz ‡ showed that a polyhedron of arbitrarily large area can be inscribed in a portion of a circular cylinder, thus proving that the ordinary definition of the length of a curve cannot be generalized directly to surfaces, the question of setting up a definition of area which would yield a suitable theory of the area of surfaces has been of great interest. Many definitions have been proposed, each one being set up to possess certain analytic or geometric properties analogous to those of the definition of length. Of these the most important are (1) those due to Lebesgue,§ Geöcze,¶ ‖ Banach,** and Peano,†† as extended by Geöcze,¶ (2) the "two dimensional measures" of a space set given by Caratheodory,‡‡ Jansen,§§ ‖‖ and Gross ¶¶ ‖‖ which can be modified to apply to "path" surfaces, and (3) other definitions, due to Minkowski,*** Young,††† Nalli and Andreoli,‡‡‡ and others,‖ which apply

* Presented to the American Mathematical Society, October 29, 1932. Literature will be cited as follows:

† National Research Fellow.

‡ H. A. Schwarz, "Gesammelte Abhandlungen," Berlin, 1891, vol. 1, pp. 309, 369.

§ H. Lebesgue, "Intégrale, longueur, aire" (Dissertation), *Annali di Matematica*, ser. 3, vol. 7 (1902), pp. 298-318 particularly.

¶ Z. de Geöcze, (a) "Über die rektifizierbare Fläche" (Hungarian), *Mathematikai és Termeszettudomanyi Ertesitö*, vol. 34 (1916), pp. 337-354; (b) "Über die Peano'sche Definition des Flächenmasses" (Hungarian), *ibid.*, vol. 35 (1917), pp. 325-360.

‖ See "Further Literature" at the end of Part I.

** S. Banach, "Sur les lignes rectifiables, etc.," *Fundamenta Mathematicae*, vol. 7 (1925), pp. 225-236.

†† G. Peano, "Sulla definizione dell' area di una superficie," *Atti della Reale Accademia dei Lincei*, ser. 4, vol. 6 (1890), pp. 54-57.

‡‡ C. Caratheodory, "Über das linear Mass von Punktmengen," *Göttingen Nachrichten*, vol. for 1914, pp. 404-426.

§§ O. Jansen, "Über einige stetige Kurven, über Bogenlänge, linearen Inhalt, und Flächeninhalt" (Dissertation), Konigsberg, 1907.

¶¶ W. Gross, "Über das Flächenmass von Punktmengen," *Monatshefte für Mathematik und Physik*, vol. 29 (1918), pp. 145-176.

‖‖ J. Schauder, "The theory of surface measure" (Thesis), *Fundamenta Mathematicae*, vol. 8 (1926), pp. 1-48.

*** H. Minkowski, "Über die Begriffe Länge, etc.," *Jahresbericht der Deutschen Mathematiker Vereinigung*, vol. 9 (1901), pp. 115-121.

††† W. H. Young, "On the area of surfaces," *Proceedings of the Royal Society of London*, ser. A, vol. 96 (1920), pp. 71-81.

‡‡‡ P. Nalli and G. Andreoli, "Sull' area di una superficie, etc.," *Atti della Reale Accademia dei Lincei*, ser. 6, vol. 5 (1927), pp. 963-966.

only to certain classes of surfaces. Unfortunately there exist examples of surfaces which are even $1-1$ continuous images of the closed unit square and for which these definitions do not all coincide.

Of those definitions applying to all surfaces, only those of Lebesgue, Geöcze, and Peano-Geöcze possess the property of *lower-semicontinuity* over the entire field of surfaces. This property is a very important property in the calculus of variations and is possessed by the area integral over the class of parametric representations in each of which all the representing functions possess continuous partial derivatives of the first order. For this reason it would appear that these definitions should be the most useful in analysis.

The first important work on these functionals was done by Geöcze * † in a series of papers appearing between 1908 and 1917. He was the first to give an analytic necessary and sufficient condition that a surface $z=f(x,y)$ possess finite Lebesgue area ‡ and proved also that $L(S)=G(S)=P(S)$ and that all are given by the classical double integral (sense of Lebesgue) when S is represented parametrically by functions satisfying a uniform Lipschitz condition, $L(S)$, $G(S)$, and $P(S)$ denoting respectively the Lebesgue, Geöcze, and Peano-Geöcze areas of S. Tonelli § took a great step forward when he defined absolutely continuous functions and functions of bounded variation of two variables so that the known theorems on the length of curves $y=f(x)$ generalize verbatim to theorems on the Lebesgue area of surfaces $z=f(x,y)$, his beautiful and new contribution being that a necessary and sufficient condition for $L(S)$ to be given by $\iint \sqrt{1+f_x{}^2+f_y{}^2}\,dxdy$ is that $f(x,y)$ be absolutely continuous in his sense. These absolutely continuous functions have proved invaluable in the present paper and in the recent independent work of McShane.¶ ‖ Radó ** †† has contributed greatly

* Z. de Geöcze, *loc. cit.*

† See "Further Literature" at the end of Part I.

‡ Z. de Geöcze, "Die notwendigen und hinreichenden Bedingungen für einen endlichen Flächeninhalt eines Flächenstückes," *Mathematikai és Physikai Lapok*, vol. 25 (1916), pp. 61-81.

§ L. Tonelli, "Sulla quadratura delle superficie," *Atti della Reale Accademia dei Lincei*, ser. 6, vol. 3 (1926), pp. 357-362, 445-450, 633-638, 714-719.

¶ E. J. McShane, "Integrals over surfaces in parametric form," *Annals of Mathematics*, vol. 34 (1933).

‖ E. J. McShane, "Parametrizations of saddle surfaces, with application to the problem of Plateau," *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 716-733.

** T. Radó, "Sur l'aire de surfaces courbes," *Acta Szeged*, vol. 3 (1927), pp. 131-169.

†† T. Radó, "Über das Flächenmass rektifizierbarer Flächen," *Mathematische Annalen*, vol. 100 (1928), pp. 445-479.

to the theory of area by first greatly simplifying the work of Geöcze, second showing that $G(S) = L(S)$ for surfaces $z = f(x, y)$, and finally giving a simple formula for $L(S)$ for such surfaces without restriction on $f(x, y)$. Saks * has shown that $L(S) = B(S)$, the Banach area of S for surfaces $z = f(x, y)$, and has recently † given an example of such a surface where $f(x, y)$ is absolutely continuous in Tonelli's sense but which nevertheless possesses a tangent plane at no point. Although Young's definition does not apply to all surfaces, he found a very general class to which it did apply and a general class of parametric representations of such surfaces for which his area was given by the usual double integral. A number of results very similar to certain ones in Sections 1, 4, and 6 of the present paper have been proved independently in a recent paper by McShane.‡

Previous to Tonelli, Evans,§ in connection with researches in potential theory, defined the notion of a "potential function of its generalized derivatives" and demonstrated many properties of these functions. Bray ¶ demonstrated a theorem similar to Lemma 4, § 4 of the present paper in which $x(u, v)$ and $y(u, v)$ were continuous potential functions of their generalized derivatives with x_u^2, y_u^2, x_v^2, and y_v^2 summable. Evans ‖ has recently shown that the notion of a continuous potential function of its generalized derivatives is identical with Tonelli's notion of an absolutely continuous function. This result with Evans' theorems on the potential functions adds a great deal to the theory of Tonelli's absolutely continuous functions.

The following is a brief summary of the results of the present paper: (1) The first section presents a systematic development of a number of theorems concerning functions of two variables which are absolutely continuous or of bounded variation in the sense of Tonelli. Most of these theorems are known although the last two have not, to the knowledge of the author, appeared in the literature, except that McShane ** has proved a theorem re-

* S. Saks, "Sur l' aire des surfaces $z = f(x, y)$," *Acta Szeged*, vol. 3 (1927), pp. 170-176.

† S. Saks, "On the surfaces without tangent planes," *Annals of Mathematics*, vol. 34 (1933), pp. 114-124.

‡ E. J. McShane, "Integrals over surfaces in parametric form," *Annals of Mathematics*, vol. 34 (1933).

§ G. C. Evans, "Fundamental points of potential theory," *Rice Institute Pamphlets*, vol. 7, no. 4 (1920), pp. 252-329.

¶ H. E. Bray, "Proof of a formula for an area," *Bulletin of the American Mathematical Society*, vol. 29 (1923), pp. 264-270.

‖ G. C. Evans, "Complements of potential theory (II)," *American Journal of Mathematics*, vol. 55 (1933), pp. 29-49.

** E. J. McShane, "Integrals over surfaces in parametric form," *Annals of Mathematics*, vol. 34 (1933).

sembling Theorem 7 but requiring uniform convergence of the functions involved. The proofs of the known theorems of this section and Section 3 are included first because they are simpler than the existing proofs and second because they generalize immediately to functions of n variables. (2) The second section recalls the definitions of $L(S)$ and the Fréchet distance of two surfaces and the elementary properties of these definitions. It then gives an exceedingly simple treatment of the functional $G(S)$. (3) The third section treats the well known theorems on the area of surfaces $z = f(x, y)$. (4) In the fourth section surfaces of "class L" are defined and for such surfaces it is shown that $L(S) = G(S)$ and that both are finite and given by the classical double integral. This class is the most general* class of surfaces yet defined for which the area in any sense is given by the integral formula; there exist examples of surfaces of this class for which $B(S) = +\infty$. (5) Section five indicates an extension of all the previous results to functions of n variables and n-dimensional manifolds. (6) Section six applies certain results of the first to obtain, in a very simple way, three theorems about "generalized conformal" representations of surfaces, these forming a very interesting sub-class of representations of class L.

Part II of this paper, which will appear in a forthcoming issue of this Journal, comprises sections seven and eight, which give a generalization of Green's formula (for space) and Stokes' formula to situations where the surfaces involved are of class L.

1. *On two classes of functions of two variables.* Following the idea of Tonelli, we define functions of two variables which are absolutely continuous or of bounded variation as follows:

Definition 1. A function, $f(x, y)$ defined in a region, R, is said to be of *bounded variation in the sense of Tonelli* (*B. V. T.*), if

$$V^{(y)}_{R_x(X)}[f(X, y)] < \mu(X), \quad V^{(x)}_{R_y(Y)}[f(x, Y)] < \mu(Y), \quad \int_{-\infty}^{\infty} \mu(s)\,ds < \infty,$$

where $V^{(y)}_{R_x(X)}[f(X, y)]$, for instance, denotes the variation of $f(X, y)$, considered as a function of y alone, over the set, $R_x(X)$, which the line $x = X$ has in common with R (being zero if $R_x(X)$ is null, and being finite or $+\infty$ otherwise; since $\mu(s)$ is summable, the variations must be finite for almost all values of the large variables). If $f(x, y)$ is continuous, these variations are lower semi-continuous and hence themselves summable.

---

* This class is more general than the one described in an abstract by the author having the title of the present paper and presented to the society on October 29, 1932. In the final draft of the present paper, the generalization to the present case was seen to present no new difficulties. This class is equivalent to that defined by McShane.

Definition 2. A function, $f(x,y)$, defined in R is said to be ***absolutely continuous in the sense of Tonelli*** $(A.C.T.)$ if

(i) $f(x,y)$ is continuous and B. V. T.

(ii) for almost all values of X, $f(X,y)$ is absolutely continuous in y in each interval of $R_x(X)$, and for almost all values of Y, $f(x,Y)$ is absolutely continuous in x in each interval of $R_y(Y)$.

We shall hereafter assume that our functions are defined on the square, $Q: 0 \leq x \leq 1,\ 0 \leq y \leq 1$, this being sufficiently general for the subsequent results. However, it is easy to see that all the theorems of this section may be extended to the case where the functions are defined in a general open region R.

The following lemma * does much to simplify the proof of the theorems to follow.

LEMMA 1. *Let* $|f(x,y)|^p$, $p \geq 1$, *be summable over* Q *and define*

$$f^{(h,k)}(x,y) = f_{h,k}(x,y) = (1/hk)\int_x^{x+h}\int_y^{y+k} f(\xi,\eta)\,d\xi d\eta, \quad h>0,\ k>0.$$

Then

$$\text{(i)}\quad \int_0^{1-h}\int_0^{1-k} |f_{h,k}(x,y)|^p\,dxdy \leq \int_0^1\int_0^1 |f(x,y)|^p\,dxdy,$$

$$\text{(ii)}\quad \lim_{\substack{h\to 0\\ k\to 0}} \int_0^{1-\alpha}\int_0^{1-\beta} |f_{h,k}(x,y) - f(x,y)|^p\,dxdy = 0.$$

Proof. The proof of (i) follows easily from the Hölder inequality as follows:

$$\int_0^{1-h}\int_0^{1-k} \left|(1/hk)\int_0^h\int_0^k f(x+\xi, y+\eta)\,d\xi d\eta\right|^p dxdy$$

$$\leq (1/hk)\int_0^{1-h}\int_0^{1-k}\left\{\int_0^h\int_0^k |f(x+\xi,y+\eta)|^p\,d\xi d\eta\right\}dxdy$$

$$= (1/hk)\int_0^h\int_0^k\left\{\int_\xi^{1-h+\xi}\int_\eta^{1-k+\eta}|f(x,y)|^p\,dxdy\right\}d\xi d\eta$$

$$\leq (1/hk)\int_0^h\int_0^k\left\{\int_0^1\int_0^1 |f(x,y)|^p\,dxdy\right\}d\xi d\eta.$$

To prove (ii), let $\{f_n(x,y)\}$ be any sequence of continuous functions so that

$$\lim_{n\to\infty}\int_0^1\int_0^1 |f(x,y) - f_n(x,y)|^p\,dxdy = 0.$$

* These "mean value functions" were used by Bray and Radó in their work on area, *loc. cit.*

Then

$$\int_0^{1-\alpha}\int_0^{1-\beta}|f(x,y)-f_{h,k}(x,y)|^p\,dxdy \leq 4^{p-1}\Big\{\int_0^{1-\alpha}\int_0^{1-\beta}|f-f_n|^p\,dxdy$$
$$+\int_0^{1-\alpha}\int_0^{1-\beta}|f_n-f_n^{(h,k)}|^p\,dxdy+\int_0^{1-\alpha}\int_0^{1-\beta}|f_n^{(h,k)}-f^{(h,k)}|^p\,dxdy\Big\}$$
$$\leq 4^{p-1}\{2\int_0^1\int_0^1|f-f_n|^p\,dxdy+\int_0^{1-\alpha}\int_0^{1-\beta}|f_n-f_n^{(h,k)}|^p\,dxdy\},$$

using (i). Since (ii) is certainly true for the continuous functions f_n, we may, for any $\epsilon > 0$, first choose n so large that the first term is $< \epsilon/2$ and then, for that n, choose h and k so small that the second term is also $< \epsilon/2$. From this the result follows.

THEOREM 1. *If $f(x,y)$ is continuous and B. V. T., then $\partial f/\partial x$ and $\partial f/\partial y$ exist almost everywhere and are summable.*

Proof. This follows because (1) all the Dini partial derivatives are measurable; (2) for almost all X, all the partials with respect to y are equal except on a set of measure zero, and for almost all Y, all the partials with respect to x are equal except on a set of measure zero; and

$$(3)\qquad \begin{aligned}&\int_0^1\Big\{\int_0^1|f_x|\,dx\Big\}dy \leqq \int_0^1 V_0^{(x)1}[f(x,Y)]dY < \infty;\\ &\int_0^1\Big\{\int_0^1|f_y|\,dy\Big\}dx \leqq \int_0^{1-h} V_0^{(y)1}[f(X,y)]dX < \infty.\end{aligned}$$

THEOREM 2. *If $f(x,y)$ is A. C. T., then*

$$\frac{\partial f^{(h,k)}}{\partial x}=\frac{1}{hk}\int_x^{x+h}\int_y^{y+k}\frac{\partial f}{\partial \xi}\,d\xi d\eta;\quad \frac{\partial f^{(h,k)}}{\partial y}=\frac{1}{hk}\int_x^{x+h}\int_y^{y+k}\frac{\partial f}{\partial \eta}\,d\xi d\eta.$$

Proof. It is sufficient to prove the first of these. Define

$$g_h(x,y)=g^{(h)}(x,y)=(1/h)\int_x^{x+h}f(\xi,y)\,d\xi.$$

Then

$$\frac{\partial g^{(h)}}{\partial x}=\frac{f(x+h,y)-f(x,y)}{h};\quad f_{h,k}(x,y)$$
$$=\frac{1}{k}\int_y^{y+k}g_h(x,\eta)\,d\eta;\quad \frac{\partial f^{(h,k)}}{\partial x}=\frac{1}{k}\int_y^{y+k}\frac{\partial g^{(h)}(x,\eta)}{\partial x}\,d\eta.$$

Now, for almost all y,

$$\frac{\partial g^{(h)}}{\partial x}=\frac{1}{h}\int_x^{x+h}\frac{\partial f(\xi,y)}{\partial \xi}\,d\xi.$$

Hence, since f_x is summable,

$$\frac{\partial f^{(h,k)}}{\partial x}=\frac{1}{k}\int_y^{y+k}\left\{\frac{1}{h}\int_x^{x+h}\frac{\partial f(\xi,\eta)}{\partial \xi}\,d\xi\right\}d\eta=\frac{1}{hk}\int_x^{x+h}\int_y^{y+k}\frac{\partial f(\xi,\eta)}{\partial\xi}\,d\xi d\eta.$$

THEOREM 3. *If $f(x,y)$ is A.C.T. and $|f_x|^p$, $|f_y|^q$ $(p,q\geqq 1)$, are summable, then*

(i) $$\int_0^{1-h}\int_0^{1-k}|f_x^{(h,k)}|^p\,dxdy\leqq\int_0^1\int_0^1|f_x|^p\,dxdy;$$

$$\int_0^{1-h}\int_0^{1-k}|f_y^{(h,k)}|^q\,dxdy\leqq\int_0^1\int_0^1|f_y|^q\,dxdy;$$

(ii) $$\lim_{\substack{h\to0\\k\to0}}\int_0^{1-\alpha}\int_0^{1-\beta}[\,|f_x^{(h,k)}-f_x|^p+|f_y^{(h,k)}-f_y|^q\,]\,dxdy=0.$$

Proof. This follows immediately from Lemma 1 and Theorem 2.

We shall now define the "x and y variations," $V_Q^{(x)}(f)$ and $V_Q^{(y)}(f)$, of $f(x,y)$ over Q in a manner analogous to the way in which the variation of a function of a single variable is defined. We shall then see that it is possible, for continuous functions, to define the two above classes of functions in a manner precisely analogous to the way in which they are defined in the case of one variable.*

Definition 3. Given an interval, $I: a\leqq x\leqq b,\ c\leqq y\leqq d$, we define the functions

$$\alpha(I)=\int_c^d[f(b,y)-f(a,y)]dy;\qquad \beta(I)=\int_a^b[f(x,d)-f(x,c)]dx.$$

Definition 4. We define the *x and y variations* $V^{(x)b,d}_{a,c}(f)$ *and* $V^{(y)b,d}_{a,c}$ of f over the rectangle $(a,b;\,c,d)$ as the variation of the set functions $\alpha(I)$ and $\beta(I)$ respectively over $(a,b;\,c,d)$, i. e. the least upper bound, for all subdivisions of $(a,b;\,c,d)$, by lines parallel to the axes, into intervals $I_1,\cdots,I_n$, of $\sum_{i=1}^n|\alpha(I_i)|$ and $\sum_{i=1}^n|\beta(I_i)|$, respectively.

LEMMA 2. *If $f(x)$ is continuous on (a,b), then*

(i) $$\int_a^{b-h}|f'_h(x)|\,dx\leqq V_a^{(x)b}[f(x)];$$

(ii) $$V_a^{(x)b}[f(x)]=\lim_{h\to0}\int_a^{b-h}|f'_h(x)|\,dx;\quad f_h(x)=(1/h)\int_x^{x+h}f(\xi)d\xi.$$

Proof. Let C_n, $C_n: y=f_n(x)$, be a sequence of polygons inscribed in

---

* Cf. Evans, *loc. cit.* (1933).

the curve $y=f(x)$ such that the length of each side is less than δ_n, $\lim_{n\to\infty}\delta_n=0$. Then

$$(\alpha)\qquad \int_a^{b-h}|f_n^{(h)\prime}(x)|\,dx \leq \int_a^b |f'_n(x)|\,dx = V_a^{(x)b}(f_n) \leq V_a^{(x)b}(f);$$

$$(\beta)\qquad \lim_{n\to\infty}\int_a^{b-h}|f_n^{(h)\prime}(x)|\,dx = \int_a^{b-h}|f'_h(x)|\,dx,\quad h>0;$$

$$(\gamma)\qquad V_a^{(x)b}(f) \leq \varliminf_{m\to\infty} V_a^{(x)b-\lambda_m}(g_m),\quad |f(x)-g_m(x)|<\mu_m,$$

$$\lim_{m\to\infty}[\,|\lambda_m|+|\mu_m|\,]=0.$$

(α) follows from Lemma 1 for one variable, since each f_n is absolutely continuous; (β) from the uniform convergence, for $h>0$, of $f_n^{(h)\prime}(x)$ to $f'_h(x)$; and (γ) directly from the definition of variation. From these three statements, the lemma follows.

THEOREM 4. *If $f(x,y)$ is continuous, and $(a,b;c,d)$ is a rectangle in Q,*

$$V^{(x)b,d}_{\ a,c}[f(x,y)] = \lim_{\substack{h\to0\\k\to0}}\int_a^{b-h}\int_c^{d-k}|f_x^{(h,k)}|\,dxdy = \int_c^d V_a^{(x)b}[f(x,Y)]dY;$$

$$V^{(y)b,d}_{\ a,c}[f(x,y)] = \lim_{\substack{h\to0\\k\to0}}\int_a^{b-h}\int_c^{d-k}|f_y^{(h,k)}|\,dxdy = \int_a^b V_c^{(y)d}[f(X,y)]dX.$$

Proof. In the first place, suppose I, $I:\alpha\leq x\leq\beta$, $\gamma\leq x\leq\delta$, is the sum of I_1, $I_1:\alpha\leq x\leq\beta$, $\gamma\leq x\leq\kappa$, and $I_2:\alpha\leq x\leq\beta$, $\kappa\leq y\leq\delta$. Then $|\alpha(I)|\leqq|\alpha(I_1)|+\alpha(I_2)|$. Thus $V^{(x)b,d}_{\ a,c}(f)$ is the upper limit, for some sequence of divisions of (a,b), $a=x_0<x_1<\cdots<x_m=b$, of

$$\sum_{j=1}^m\int_c^d|f(x_j,y)-f(x_{j-1},y)|\,dy = \int_c^d\phi_m(y)\,dy,$$

where $\phi_m(y)=\sum_{j=1}^m|f(x_j,y)-f(x_{j-1},y)|$. Since $\{\phi_m(y)\}$ converges to $V_a^{(x)b}[f(x,y)]$ for almost all y, and since $0\leq\phi_m(y)\leq V_a^{(x)b}[f(x,y)]$, the latter being a summable function of y, it is easy to see that

$$V^{(x)b,d}_{\ a,c}(f) = \int_c^d V_a^{(x)b}[f(x,Y)]dY.$$

Now define

$$g_h(x,y) = (1/h)\int_x^{x+h}f(\xi,y)\,d\xi = g^{(h)}(x,y).$$

Then $\partial g^{(h)}/\partial x$ is continuous and

$$\frac{\partial f^{(h,k)}}{\partial x} = \frac{1}{k}\int_y^{y+k}\frac{\partial g^{(h)}(x,\eta)}{\partial x}\,d\eta.$$

Now we know that

$$V^{(x)b-h,d}_{a,c}[g_h(x,y)] = \int_c^d V_a^{(x)b-h}[g_h(x,Y)]dY = \int_a^{b-h}\int_c^d | g_x^{(h)}(x,y) | \, dxdy$$

$$\leq \int_c^d V_a^{(x)b}[f(x,Y)]dY = V^{(x)b,d}_{a,c}(f).$$

$$\therefore V^{(x)b-h,d-k}_{a,c}(f_{h,k}) = \int_a^{b-h}\int_c^{d-k} | f_x^{(h,k)} | \, dxdy$$

$$\leq \int_a^{b-h}\int_c^d | g_x^{(h)} | \, dxdy \leq V^{(x)b,d}_{a,c}(f).$$

On the other hand it is clear that

$$V^{(x)b,d}_{a,c}(f) \leq \lim_{n\to\infty} V^{(x)b-\alpha_n,d-\beta_n}_{a,c}(f_n), \quad | f - f_n | < \gamma_n,$$

$$\lim_{n\to\infty} [\, | \alpha_n | + | \beta_n | + | \gamma_n | \,] = 0,$$

in the same way that it is shown in one dimension. Thus the result follows for the x variation and a similar proof holds for the y variation.

THEOREM 5. *A necessary and sufficient condition that the continuous function $f(x,y)$ be B. V. T. is that its x and y variations be both finite. A necessary and sufficient condition for $f(x,y)$ to be A. C. T. is that it be continuous and* (a) *the set functions $\alpha(I)$ and $\beta(I)$ be absolutely continuous or* (b) *we have*

$$V^{(x)b,d}_{a,c}(f) = \int_a^b\int_c^d | f_x | \, dxdy, \quad V^{(y)b,d}_{a,c}(f) = \int_a^b\int_c^d | f_y | \, dxdy,$$

for every interval, $a \leq x \leq b$, $c \leq y \leq d$, in Q.

Proof. The first statement follows immediately from the preceding theorems. The "necessary part" of the second statement also follows from that and the preceding theorems. If f is B. V. T. and the formulas (b) hold, then it is clear that $\alpha(I)$ and $\beta(I)$ are absolutely continuous set functions.

Hence assume $\alpha(I)$ and $\beta(I)$ are absolutely continuous and $f(x,y)$ continuous. Then $f(x,y)$ is B. V. T. Let us consider the interval function $\alpha(I)$. We know (1) that it has a derivative almost everywhere and (2) its variation is the integral of the absolute value of this derivative. Thus, for almost every (x,y),

$$\lim_{\substack{h\to 0\\ k\to 0}} (1/hk)\int_y^{y+k}[f(x+h,\eta) - f(x,\eta)]d\eta = \lim_{h\to 0}\frac{f(x+h,y) - f(x,y)}{h} = f_x,$$

since it is clear that f_x exists almost everywhere that the above limit as $h \to 0$ does. Furthermore we know that

$$V^{(x)b,d}_{a,c}(f) = \int_a^b \int_c^d |f_x|\, dxdy = \int_c^d V_a^{(x)b}[f(x,Y)]dY.$$

Hence, for almost every Y, we must have that

$$\int_a^b |f_x|\, dx = V_a^{(x)b}[f(x,Y)]$$

since for almost every Y,

$$\int_a^b |f_x|\, dx \leq V_a^{(x)b}[f(x,Y)]$$

and f is B. V. T. Thus $f(x,y)$ is absolutely continuous in x for almost all y. The corresponding results with the rôles of x and y interchanged are proved in the same way starting with $\beta(I)$.

THEOREM 6. *Let $f(x,y)$ be continuous and $h(t)$, $k(t)$ positive functions approaching zero with t. Suppose*

$$(1.1) \quad \int_0^{1-h(t)} \int_0^{1-k(t)} [\,|f_x^{(h(t),k(t))}|^p + |f_y^{(h(t),k(t))}|^q\,]\, dxdy < M, \quad p, q > 1$$

for every $t > 0$. Then $f(x,y)$ is A. C. T. and $|f_x|^p$ and $|f_y|^q$ are summable.

Proof. First of all, let us suppose that we have two sequences $\{\alpha_i\}$, $\{\beta_i\}$ where $\alpha_i \geq 0$ and $\beta_i > 0$. Then if $\lambda > 0$, it is easy to see that

$$\sum_{i=1}^{\infty} \frac{\alpha_i^{1+\lambda}}{\beta_i^{\lambda}} \geq \frac{(\sum_{i=1}^{\infty} \alpha_i)^{1+\lambda}}{(\sum_{i=1}^{\infty} \beta_i)^{\lambda}}, \quad \text{if} \quad \sum_{i=1}^{\infty} \beta_i < \infty,$$

for it is easily verified when there are just two non-zero terms in the series on the left and may then be proved in general by induction and a simple limit process (the value $+\infty$ being allowed).

Now let us suppose that $f(x,y)$ is not A. C. T. Then there exists an $\epsilon > 0$ and sequences, $\{I_{m,n}\}$, of non-overlapping intervals such that

$$\sum_{n=1}^{\infty} \text{meas}(I_{m,n}) = \delta_m, \quad \sum_{n=1}^{\infty} [\,|\alpha(I_{m,n})| + |\beta(I_{m,n})|\,] > 2\epsilon; \quad m = 1, 2, \cdots, \lim_{m\to\infty} \delta_m = 0;$$

so that either $\sum_{n=1}^{\infty} |\alpha(I_{m,n})| > \epsilon$ or $\sum_{n=1}^{\infty} |\beta(I_{m,n})| > \epsilon$. Assume the former, for instance. Then clearly

$$\varliminf_{t\to 0}\int_0^{1-h(t)}\int_0^{1-k(t)} |f_x{}^{(h(t),k(t))}|^p\,dxdy \geq \sum_{n=1}^{\infty}\varliminf_{t\to 0}\int_{a_{m,n}}^{b_{m,n}-h(t)}\int_{c_{m,n}}^{d_{m,n}-k(t)} |f_x{}^{(h(t),k(t))}|^p\,dxdy$$

$$\geq \sum_{n=1}^{\infty}\left[\varliminf_{t\to 0}\frac{\left\{\int_{a_{m,n}}^{b_{m,n}-h(t)}\int_{c_{m,n}}^{d_{m,n}-k(t)} |f_x{}^{(h(t),k(t))}|\,dxdy\right\}^p}{\{\operatorname{meas}(I_{m,n})\}^{p-1}}\right.$$

$$\geqq \sum_{n=1}^{\infty}\frac{|\alpha(I_{m,n})|^p}{[\operatorname{meas}(I_{m,n})]^{p-1}} \geq \frac{[\sum_{n=1}^{\infty}|\alpha(I_{m,n})|]^p}{[\sum_{n=1}^{\infty}\operatorname{meas}(I_{m,n})]^{p-1}} \geqq \frac{\epsilon^p}{\delta_m{}^{p-1}}$$

which may be made as large as we please by taking m large enough. This however contradicts the hypothesis (1.1). Thus $f(x,y)$ must be A. C. T.

We know that, almost everywhere

$$\lim_{t\to 0}\frac{\partial f^{(h(t),k(t))}}{\partial x}=\frac{\partial f}{\partial x};\quad \lim_{t\to 0}\left|\frac{\partial f^{(h(t),k(t))}}{\partial x}\right|^p=|f_x|^p;\quad \lim_{t\to 0}\left|\frac{\partial f^{(h(t),k(t))}}{\partial y}\right|^q=|f_y|^q,$$

by Lemma 1, since f is A. C. T. But now, by a well known theorem, $|f_x|^p$, $|f_y|^q$ are summable and

$$\iint_{Q^\alpha}|f_x|^p\,dxdy\leq\varliminf_{t\to 0}\iint_{Q^\alpha}|f_x{}^{(h(t),k(t))}|^p;$$

$$\iint_{Q^\alpha}|f_y|^q\,dxdy\leq\varliminf_{t\to 0}\iint_{Q^\alpha}|f_y{}^{(h(t),k(t))}|^q\,dxdy,\quad Q^\alpha:\ 0\leq x,\ y\leq 1-\alpha.$$

THEOREM 7. *Let $f_n(x,y)$ be A. C. T. and $f(x,y)$ continuous and suppose*

(i) $\left|\int_0^x\int_0^y[f_n(\xi,\eta)-f(\xi,\eta)]d\xi d\eta\right|<\epsilon_n,\quad \lim_{n\to\infty}\epsilon_n=0;$

(ii) $\int_0^1\int_0^1[\,|(f_n)_x|^p+|(f_n)_y|^q\,]dxdy<M,\quad p,q>1.$

Then $f(x,y)$ is A. C. T. and

$$\int_0^1\int_0^1|f_x|^p\,dxdy\leq\varliminf_{n\to\infty}\int_0^1\int_0^1|(f_n)_x|^p\,dxdy;$$

$$\int_0^1\int_0^1|f_y|^q\,dxdy\leq\varliminf_{n\to\infty}\int_0^1\int_0^1|(f_n)_y|^q\,dxdy.$$

Proof. Let $h>0$, $k>0$. Then $f_n{}^{(h,k)}(x,y)$ approaches $f^{(h,k)}(x,y)$ uniformly as to x and y. Hence, for $h'>0$, $k'>0$, $\dfrac{\partial(f_n{}^{(h,k)})^{(h',k')}}{\partial x[\partial y]}$ approaches $\dfrac{\partial(f^{(h,k)})^{(h',k')}}{\partial x[\partial y]}$ uniformly. Hence

$$\lim_{n\to\infty}\int_0^{1-h-h'}\int_0^{1-k-k'}|(f_n^{(h,k)})_x^{(h',k')}|^p\,dxdy=\int_0^{1-h-h'}\int_0^{1-k-k'}|(f^{(h,k)})_x^{(h',k')}|^p\,dxdy.$$

Now since $f(x,y)$ and therefore $f_x^{(h,k)}$ is continuous, we see that

$$\lim_{\substack{h\to 0\\ k\to 0}}\int_0^{1-h-h'}\int_0^{1-k-k'}|(f^{(h,k)})_x^{(h',k')}|^p\,dxdy=\int_0^{1-h}\int_0^{1-k}|f_x^{(h,k)}|^p\,dxdy.$$

From this it follows that

$$\int_0^{1-h}\int_0^{1-k}|f_x^{(h,k)}|^p\,dxdy\leq\varliminf_{n\to\infty}\int_0^{1-h}\int_0^{1-k}|(f_n^{(h,k)})_x|^p\,dxdy\leq M.$$

Since the rôles of x and y may be interchanged, $f(x,y)$ is seen, using Theorem 6, to be A. C. T. with $|f_x|^p$ and $|f_y|^q$ summable. Then, using Theorem 3, we see that

$$\lim_{\substack{h\to 0\\ k\to 0}}\int_0^{1-h}\int_0^{1-k}|f_x^{(h,k)}|^p\,dxdy=\int_0^1\int_0^1|f_x|^p\,dxdy;$$

$$\lim_{\substack{h\to 0\\ k\to 0}}\int_0^{1-h}\int_0^{1-k}|(f_n^{(h,k)})_x|^p\,dxdy=\int_0^1\int_0^1|(f_n)_x|^p\,dxdy.$$

From this, with the similar equations obtained by interchanging the rôles of x and y, the conclusion follows immediately.

2. *On certain definitions of area.* In this section, we shall consider two definitions of the area of a continuous surface and shall develop some of their elementary geometric properties. The first of these is the well known definition of Lebesgue * and the second is that due to Geöcze.

In order to proceed with the discussion, we shall first define the (Fréchet) distance, $\| S_1, S_2 \|$, of two surfaces S_1 and S_2.

Definition 1. Let $S_1 : x^i = x_1{}^i(u,v)$, $S_2 : x^i = X_2{}^i(s,t)$, $i=1,\cdots,N$, (u,v), (s,t) in Q and $x_1{}^i(u,v)$ and $X_2{}^i(s,t)$ continuous. Let T,

$$T : s = s(u,v),\ t = t(u,v),$$

be a $1-1$ continuous, sense preserving transformation of Q into itself. Define $x_2{}^i(u,v) = X_2{}^i[s(u,v),\ t(u,v)]$ and

$$D_T(S_1,S_2)=\max_{(u,v)\in Q}\sqrt{\sum_{i=1}^{N}[x_1{}^i(u,v)-x_2{}^i(u,v)]^2}.$$

Then $\| S_1, S_2 \|$ is the greatest lower bound of $D_T(S_1,S_2)$ for all T.

* For a systematic exposition of the elementary properties of Lebesgue area, see Radó, *loc. cit.* (*Acta Szeged*).

Definition 2. If $\| S_1, S_2 \| = \| S_2, S_1 \| = 0$, we say that $S_1 \equiv S_2$ and that *the two sets of functions represent the same surface.*

The following lemma is an immediate consequence of these definitions.

LEMMA 1. (a) $\| S_1, S_2 \| = \| S_2, S_1 \|$; (b) $\| S_1, S_3 \| \leq \| S_1, S_2 \| + \| S_2, S_3 \|$; (c) *if* $\| S_1, \bar{S}_1 \| = \| S_2, \bar{S}_2 \| = 0$, *then* $\| \bar{S}_1, \bar{S}_2 \| = \| S_1, S_2 \|$; (d) *if* $\lim_{n\to\infty} \| S, S_n \| = 0$ *and* $x^i = x^i(u, v)$ *is any representation of* S, *we can find representations,* $x^i = x_n{}^i(u, v)$, *of* S_n *so that the* $x_n{}^i(u, v)$ *converge uniformly to the* $x^i(u, v)$, $i = 1, \cdots, N$.

Definition 3. If $\lim_{n\to\infty} \| S, S_n \| = 0$, we say $\lim_{n\to\infty} S_n = S$.

Remark. The above definition and all of its properties are independent of the number of dimensions, N, and the number of parameters (u, v).

We are now in a position to define the Lebesgue area, $L(S)$, of a surface S.

Definition 4. We say that a surface, Π, is a *polyhedron* if, among its representations there is one such that Q is divided into a finite number of triangles in each of which each of the continuous representing functions is linear.

Definition 5. Given a surface S. Let $\{\Pi_n{}^{(a)}\}$ be a sequence of polyhedra approaching S. Let

$$\alpha = \varliminf_{n\to\infty} L(\Pi_n{}^{(a)})$$

the area of $\Pi_n{}^{(a)}$ being the sum of the areas of its component triangles. Then $L(S)$ is defined as the greatest lower bound of all such numbers α.

The following three theorems about Lebesgue area are well known.

THEOREM 1. *If* $\| S_1, S_2 \| = 0$, *then* $L(S_1) = L(S_2)$.

THEOREM 2. *Given any surface,* S, *there exists a sequence of polyhedra,* $\{\Pi_n\}$, *approaching* S *such that* $L(\Pi_n)$ *approaches* $L(S)$ (we admit $+\infty$ as a possible numerical value of the area of a surface when discussing any definition).

THEOREM 3. *If* $\{S_n\}$ *is a sequence of surfaces approaching* S, *then*

$$L(S) \leq \varliminf_{n\to\infty} L(S_n).$$

Let C, $C: x = x(u)$, $y = y(u)$, $0 \leq u \leq 1$, $x(0) = x(1)$, $y(0) = y(1)$, $x(u)$, $y(u)$ continuous, be a closed curve in the (x, y) plane. Let $O_{x,y}(s, t; C)$

* See, for instance, Kerékjártó, *Vorlesung über Topologie I*, zweiter Abschitt, § 2.

be the signed order * of the point $x = s$, $y = t$ with respect to the curve C if (s, t) is not on C; if (s, t) is on C, define $O_{x,y}(s, t; C) = 0$. Thus $O_{x,y}(s, t; C)$ has a definite finite integral value at every point of the plane. The following further remark is immediate:

Lemma 2. *If (s, t) is not on C and $\{C_n\}$ is a sequence of closed curves approaching C then $O_{x,y}(s, t; C_n)$ approaches $O_{x,y}(s, t; C)$.*

We shall next consider the Geöcze definition of area. It may be defined as follows:

Definition 6. Suppose S is represented on Q by the equations $x^i = x^i(u, v)$, $i = 1, \cdots, N$. Divide Q up into a finite number, $R_1, \cdots, R_n$, of Jordan regions with respective boundaries $\Gamma_1, \cdots, \Gamma_n$. Let

$$F(R_k) = \sqrt{\sum_{i=2}^{N} \sum_{j=1}^{i-1} [F_{x^i,x^j}(R_k)]^2},$$

$$F_{x^i,x^j}(R_k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | O_{x^i,x^j}(s, t; C_k^{(i,j)}) | \, ds dt \qquad (\leq + \infty),$$

where $C_k^{(i,j)}$ is the projection (obtained by suppressing the other coördinates) in the (x^i, x^j) plane of the curve C_k of S which corresponds to Γ_k. Then $G(S)$ is the least upper bound, for all such subdivisions of Q, of the sum

$$\sum_{k=1}^{n} F(R_k).$$

The remainder of this section will be devoted to the proof of several simple properties of $G(S)$ embodied in Theorems 4, 5, 6, and 7.

Theorem 4. *If* $\| S_1, S_2 \| = 0$, $G(S_1) = G(S_2)$.

Proof. This follows directly from the definitions.

Theorem 5. *If $\{S_n\}$ is a sequence of surfaces approaching S, then*

$$G(S) \leq \varliminf_{n \to \infty} G(S_n).$$

Proof. Since, for each $\epsilon > 0$, it is possible to subdivide Q into Jordan regions $R_1, \cdots, R_{n_\epsilon}$ so that

$$\sum_{i=1}^{n_\epsilon} F(R_i) > G(S) - \epsilon,$$

it is clear that it is sufficient to show that

$$F_{x^i,x^j}(R) \leq \varliminf_{n\to\infty} F^{(n)}_{x^i,x^j}(R), \qquad (i=2,\cdots,N,\ j=1,\cdots,i-1,$$

for each R in Q. But this is immediate for if (s,t) is not on C_{x^i,x^j}, $O_{x^i,x^j}(s,t;C_n^{i,j})$ approaches $O_{x^i,x^j}(s,t;C^{i,j})$, ($C^{i,j}$ being the projection of the curve C of S corresponding to Γ on the (x^i,x^j) plane) since $C_n^{i,j}$ clearly approaches $C^{i,j}$. If (s,t) is on $C^{i,j}$, then

$$|\, O_{x^i,x^j}(s,t;C^{i,j})\,| = 0 \leq \varliminf_{n\to\infty} |\, O_{x^i,x^j}(s,t;C_n^{i,j})\,|$$

THEOREM 6. *Suppose we divide S into two parts, S_1 and S_2, by means of an arc whose projection on each coördinate plane is of measure zero. Then*

$$G(S) = G(S_1) + G(S_2).$$

Proof. In the first place, suppose a Jordan region, R in Q, is divided into two Jordan regions, R_1 and R_2, by an arc γ corresponding to a curve c, whose projections on each coördinate plane are of measure zero. Let Γ be the boundary of R, Γ_k that of R_k, and C and C_k the corresponding curves of S. Now if (s,t) is not on $c^{i,j}$,

$$O_{x^i,x^j}(s,t;C^{i,j}) = O_{x^i,x^j}(s,t;C_1^{i,j}) + O_{x^i,x^j}(s,t;C_2^{i,j}).$$

Since $c^{i,j}$ is of measure zero, it follows by integration that

$$F_{x^i,x^j}(R) \leq F_{x^i,x^j}(R_1) + F_{x^i,x^j}(R_2)$$

for every pair of indices (i,j), $i\neq j$. Thus $F(R) \leq F(R_1) + F(R_2)$.

Now suppose Σ is any subdivision of Q into Jordan regions $R_1,\cdots,R_n$. Let Σ_m be a sequence of subdivisions of Q into regions $R_{1,m},\cdots,R_{n,m}$ so that $\Gamma_{i,m}$ approaches Γ_i (using our notations above). Then from the proof of Theorem 5,

$$\sum_{i=1}^{n} F(R_i) \leq \varliminf_{m\to\infty} \sum_{i=1}^{n} F(R_{i,m}).$$

Now let Γ be the simple arc in Q corresponding to the curve C which divides S into S_1 and S_2. Suppose Γ divides Q into Q_1 and Q_2. Then, let $\bar{\Sigma}$ be a subdivision so that $\sum_{i=1}^{n} F(\bar{R}_i) > G(S) - \epsilon/2$. According to the preceding paragraph, we can replace this by a subdivision into regions R_i so that each Γ_i has only a finite number of points in common with the dividing curve, Γ, and so that $\sum_{i=1}^{n} F(R_i) > G(S) - \epsilon$. Now replace Σ by the new subdivision,

Σ', consisting of all the (finite number) Jordan regions, $R'_1, \cdots, R'_{n'}$, into which Q is divided by the Γ_i and Γ. Then by repeated application of the first paragraph of the proof, we find that

$$\sum_{i=1}^{n'} F(R'_i) \geqq \sum_{i=1}^{n} F(R_i) > G(S) - \epsilon.$$

But each R'_i lies either wholly in Q_1 or wholly in Q_2. Thus it is clear that $G(S_1) + G(S_2) \geqq G(S)$. On the other hand it is obvious (since every subdivision of Q_1 plus one of Q_2 gives one of Q) that $G(S_1) + G(S_2) \leqq G(S)$; so that the theorem is completely demonstrated.

THEOREM 7. *If* Π *is a polyhedron,* $G(\Pi) = L(\Pi)$. *Thus, in general,* $G(S) \leqq L(S)$.

Proof. If Π consists of one triangle, the theorem is obvious. Using the preceding theorem, the relation may be established by induction.

3. *The area of surfaces* $z = f(x, y)$. In this section, we give an extremely simple treatment of this subject very similar to the developments of § 1.

LEMMA 1.* *Suppose that* $f^i(x)$ *is summable on* Q, $i = 1, \cdots, n$. *Then*

$$\sqrt{\sum_{i=1}^{n} \left[\int_Q f^i dx\right]^2} \leqq \int_Q \left(\sqrt{\sum_{i=1}^{n} (f^i)^2}\right) dx,$$

where the letter x *stands for* $(x^1, \cdots, x^N)$ *and* Q *is a region of the space of* $(x^1, \cdots, x^N)$.

Proof. This lemma on integrals is merely a limiting case of the inequality

$$\sqrt{\sum_{i=1}^{n} \left[\sum_{j=1}^{N} a_j{}^i\right]^2} \leqq \sum_{j=1}^{N} \sqrt{\sum_{i=1}^{n} (a_j{}^i)^2},$$

which states that the length of the sum of N vectors in n-space is not greater than the sum of the lengths of these vectors.

LEMMA 2. *If* $f(x, y)$ *is A. C. T. on* Q,

$$\int_0^{1-h} \int_0^{1-k} \sqrt{1 + [f_x{}^{(h,k)}]^2 + [f_y{}^{(h,k)}]^2}\, dxdy \leqq \int_0^1 \int_0^1 \sqrt{1 + f_x{}^2 + f_y{}^2}\, dxdy < \infty.$$

Proof. This is an immediate consequence of the above lemma and Theorem 2, § 1 as follows:

* Young, *loc. cit.*

$$\int_0^{1-h}\int_0^{1-k}\left\{\sqrt{\left[\frac{1}{hk}\int_0^h\int_0^k d\xi d\eta\right]^2+\left[\frac{1}{hk}\int_0^h\int_0^k f_\xi\,(x+\xi,y+\eta)\,d\xi d\eta\right]^2}\right.$$
$$\left.\overline{+\left[\frac{1}{hk}\int_0^h\int_0^k f_\eta(x+\xi,y+\eta)\,d\xi d\eta\right]^2}\right\}dxdy$$
$$\leq\int_0^{1-h}\int_0^{1-k}\left[\frac{1}{hk}\int_0^h\int_0^k\{\sqrt{1+[f_\xi(x+\xi,y+\eta)]^2}\right.$$
$$\left.\overline{+[f_\eta(x+\xi,y+\eta)]^2}\}d\xi d\eta\right]dxdy$$
$$=\frac{1}{hk}\int_0^h\int_0^k\left[\int_\xi^{1-h+\xi}\int_\eta^{1-k+\eta}(\sqrt{1+f_x^2+f_y^2})\,dxdy\right]d\xi d\eta$$
$$\leq\frac{1}{hk}\int_0^h\int_0^k\left[\int_0^1\int_0^1(\sqrt{1+f_x^2+f_y^2})\,dxdy\right]d\xi d\eta.$$

THEOREM 1. *Let* $S:z=f(x,y)$.* *Then*

(i) $\displaystyle\int_0^{1-h}\int_0^{1-k}[\sqrt{1+(f_x^{(h,k)})^2+(f_y^{(h,k)})^2}]\,dxdy\leq L(S)$;

(ii) $\displaystyle L(S)=\lim_{\substack{h\to 0\\ k\to 0}}\int_0^{1-h}\int_0^{1-k}[\sqrt{1+(f_x^{(h,k)})^2+(f_y^{(h,k)})^2}]\,dxdy.$

Proof. It is clear that we can choose a sequence of polyhedra, $\{\Pi_n\}$, of the form $z=f_n(x,y)$, so that $\lim_{n\to\infty}\Pi_n=S$, $\lim_{n\to\infty}L(\Pi_n)=L(S)$. Now, for each Π_n,

$$\int_0^{1-h}\int_0^{1-k}\left(\sqrt{1+\left(\frac{\partial f_n^{(h,k)}}{\partial x}\right)^2+\left(\frac{\partial f_n^{(h,k)}}{\partial y}\right)^2}\right)dxdy\leq L(\Pi_n)$$

using Lemma 2. Since $f_n(x,y)$ converges uniformly to $f(x,y)$, it is clear from the formulas for $(f_n^{(h,k)})_x$ and $(f_n^{(h,k)})_y$ that, for $h>0$, $k>0$, $(f_n^{(h,k)})_x$ and $(f_n^{(h,k)})_y$ converge uniformly (in x and y) to $f_x^{(h,k)}$ and $f_y^{(h,k)}$ respectively. Hence (i) follows immediately.

On the other hand, it is clear the surfaces $S^{(h,k)}:z=f^{(h,k)}(x,y)$, $0\leq x\leq 1-h$, $0\leq y\leq 1-k$, approach S. Therefore

$$L(S)\leq\varliminf_{\substack{h\to 0\\ k\to 0}}L(S^{(h,k)}),$$

and the theorem is completely demonstrated.

THEOREM 2.† *A necessary and sufficient condition that* $L(S)$ *be finite,* $S:z=f(x,y)$, *is that* $f(x,y)$ *be B. V. T.*

Proof. This follows from the above theorem and Theorems 4 and 5, § 1, for

* Radó, *loc. cit.* (*Acta Szeged*).

† Tonelli, *loc. cit.*

$$\int_0^{1-h}\int_0^{1-k} | f_x^{(h,k)} | \, dxdy,$$

$$\int_0^{1-h}\int_0^{1-k} | f_y^{(h,k)} | \, dxdy \leq \int_0^{1-h}\int_0^{1-k} [\sqrt{1+(f_x^{(h,k)})^2+(f_y^{(h,k)})^2}] \, dxdy$$

$$\leq 1 + \int_0^{1-h}\int_0^{1-k} [\, | f_x^{(h,k)} | + | f_y^{(h,k)} | \,] \, dxdy.$$

THEOREM 3.† *A necessary and sufficient condition that $L(S)$ be finite and given by*

$$L(S) = \iint_Q (\sqrt{1+f_x^2+f_y^2})\, dxdy,$$

is that $f(x,y)$ be A. C. T.

Proof. From Lemma 2 and Theorem 1, it is clear that the condition is sufficient.

To prove that it is also necessary, suppose that $L(S)$ is finite and given by the above formula. Let (x_0, y_0) be a point of Q and let Q be divided into the four rectangles $(0, x_0;\ 0, y_0)$, $(0, x_0;\ y_0, 1)$, etc., by the lines $x = x_0$ and $y = y_0$. It is easy to see by studying the formula of Theorem 1 and the formulas in § 1, Theorem 4, that $L(S_{0,0}^{x_0,y_0})$ is a continuous function of (x_0, y_0), $S_{0,0}^{x_0,y_0}$ being the part of S above the rectangle $(0, x_0;\ 0, y_0)$. Hence

$$L(S_{0,0}^{x_0,y_0}) + L(S_{0,y_0}^{x_0,1}) + L(S_{x_0,0}^{1,y_0}) + L(S_{x_0,y_0}^{1,1}) = L(S).$$

Furthermore, since $f(x,y)$ is B. V. T. it is clear that the set function

$$\phi(I) = \sqrt{[\alpha(I)]^2 + [\beta(I)]^2 + [\text{meas}\, I]^2}$$

is of bounded variation and thus has a derivative equal almost everywhere to $\sqrt{1+f_x^2+f_y^2}$, since the derivatives of $\alpha(I)$ and $\beta(I)$ are f_x and f_y, respectively, almost everywhere (see the proof of Theorem 5, § 1). Also the variation of $\phi(I)$ over $(a, b;\ c, d)$ is greater or equal to the integral over $(a, b;\ c, d)$ of this derivative. Combining all these facts, we see that

$$L(S_{0,0}^{x_0,y_0}) = \int_0^{x_0}\int_0^{y_0} (\sqrt{1+f_x^2+f_y^2})\, dxdy;$$

$$L(S_{a,c}^{b,d}) = \int_a^b\int_c^d (\sqrt{1+f_x^2+f_y^2})\, dxdy.$$

Using this fact, Theorem 1 and Theorem 4, § 1, we see that

$$\alpha\binom{b,d}{a,c} \leq V^{(x)}{}_{a,c}^{b,d}(f) = \lim_{\substack{h\to 0\\ k\to 0}} \int_a^{b-h}\int_c^{d-h} | f_x^{(h,k)} | \, dxdy \leq \int_a^b\int_c^d (\sqrt{1+f_x^2+f_y^2})\, dxdy$$

and that $\beta(I)$ satisfies a similar inequality. Thus $\alpha(I)$ and $\beta(I_1)$ are absolutely continuous set functions and $f(x,y)$ is A. C. T.

4. *The area of surfaces "of class L."* In this section, we extend the range of applicability of the classical formula for the area of a surface to a very general class of surfaces.

Definition 1. We say that a surface S, $S: x^i = x^i(u,v)$, $i = 1, \cdots, N$, or more properly, the given parametric representation, is of *class* L if

(i) the $x^i(u,v)$ are all A. C. T.,

$$\text{(ii)} \quad \lim_{h \to 0} \iint_Q \left| \frac{\partial(\bar{x}_h^i, \bar{x}_h^j)}{\partial(u,v)} - \frac{\partial(x^i, x^j)}{\partial(u,v)} \right| dudv = 0,$$

$$\bar{S}_h : x^i = \bar{x}_h^i(u,v), \quad \bar{x}_h^i(u,v) = x^i_{h,h}[u(1-h), v(1-h)],$$

In (ii), $x^i_{h,h}(u,v)$ is the usual mean value function defined for $0 \leq u$, $v \leq 1 - h$.

Definition 2. We shall sometimes speak of a "flat surface," $S: x = x(u,v)$, $y = y(u,v)$, as a *transformation* and shall say that a transformation is of class L if the corresponding (flat) surface is.

The following conditions define a convenient subclass of surfaces of class L (as is easily verified, using Lemma 1 and Theorems 1 and 2 of § 1, and the Hölder inequality)

(i) $x^i(u,v)$ are A. C. T., $i = 1, \cdots, N$,

(ii) $|x_u^i|^p$, $|x_v^i|^q$ are summable, $i = 1, \cdots, N$, $p, q \geq 1$, $1/p + 1/q \leq 1$,

where we include the case where one of p and q is unity and the other infinite; if $p = \infty$, $q = 1$, for instance, we interpret (ii) to mean

(ii′) $\qquad |x_u^i| < M$, $|x_v^i|$ summable, $i = 1, \cdots, N$.

Surfaces $z = f(x,y)$ with $f(x,y)$ A. C. T. are easily seen to be of class L although they do not come under the above head.

The first three of the following lemmas are well known and require no proof. The fourth is essentially new, although, as was noted in the introduction, Bray * has proved a similar theorem.

LEMMA 1. *Let* C, $C: x = x(t)$, $y = y(t)$, $x(0) = x(1)$, $y(0) = y(1)$,

* Bray, *loc. cit.*

be a closed rectifiable curve. Let $\{C_n\}$, $C_n : x = x_n(t)$, $y = y_n(t)$, *be a sequence of curves approaching* C *such that* $l(C_n) < M$. *Then*

$$\lim_{n\to\infty} \int_0^1 x_n(t)\,dy_n(t) = \int_0^1 x(t)\,dy(t).$$

LEMMA 2. *If* C, $C : x = x(t)$, $y = y(t)$, x *and* y *absolutely continuous, is a closed curve and* (s, t) *is not on* C, *then*

$$O_{x,y}(s, t; C) = \frac{1}{2\pi}\int_0^1 \frac{[x(u) - s]y'(u) - [y(u) - s]x'(u)}{[x(u) - s]^2 + [y(u) - t]^2}\,du.$$

LEMMA 3. *Given a closed curve* $C : x = x(t)$, $y = y(t)$, x *and* y *absolutely continuous. Then*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} O_{x,y}(s, t; C)\,ds\,dt = \int_0^1 x(u)y'(u)\,du.$$

Definition 3. A summable function, $f(x)$, is said to be *metrically continuous* at $x = x_0$ if

$$\lim_{h\to 0} \frac{1}{h}\int_{x'_0}^{x'_0+h} f(\xi)\,d\xi = f(x_0), \qquad x_0 - h \leq x'_0 \leq x_0.$$

Remark 1. This definition is independent of the number of variables, if we replace the h by a system $(h^1, \cdots, h^n)$, the integral by a multiple integral, and the limit by a multiple limit.

Remark 2. A summable function is metrically continuous almost everywhere in its region of definition.

LEMMA 4. *Let* $T : x = x(u, v)$, $y = y(u, v)$ *be a transformation of class* L. *Suppose:*

(1) $x(u, v)$ *and* $y(u, v)$ *are absolutely continuous around the boundary of the rectangle,* $(a, b;\ c, d)$;

(2) $V_0^{(u)1}[y(u, V)]$ *is metrically continuous in* V *for* $V = c$ *and* $V = d$;

(3) $V_0^{(v)1}[y(U, v)]$ *is metrically continuous in* U *for* $U = a$ *and* $U = b$. *Then*

$$\int_a^b\int_c^d \frac{\partial(x, y)}{\partial(u, v)}\,du\,dv = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} O_{x,y}(s, t; C)\,ds\,dt = \int_C x\,dy, \tag{4.1}$$

where C *is the closed rectifiable image of the boundary of* $(a, b;\ c, d)$.

Proof. The equality between the last two members of (4.1) follows from the preceding lemma. Since T is of class L,

$$\lim_{h\to 0}\int_a^b\int_c^d \frac{\partial(\bar{x}^{(h)}, \bar{y}^h)}{\partial(u,v)}\,dudv = \int_a^b\int_c^d \frac{\partial(x,y)}{\partial(u,v)}\,dudv.$$

Since the equality between the first and last members of (4. 1) holds for every $h > 0$, we need merely to prove that

$$\lim_{h\to 0}\int_{\bar{C}^{(h)}} \bar{x}^h d\bar{y}^h = \int_C x dy.$$

To do this, it is sufficient to show that $\int_{\bar{C}^h} | d\bar{y}^h |$ is bounded, independently of h, since $\bar{x}^h(t)$ and $\bar{y}^{(h)}(t)$ approach $x(t)$ *and* $y(t)$ uniformly (t on rectangle). Now

$$\int_{\bar{C}^h} | d\bar{y}^h | = \int_a^b [| \bar{y}_u{}^{(h)}(u,d)| + | \bar{y}_u{}^h(u,c)|]du + \int_c^d [| \bar{y}_v{}^{(h)}(b,v)| + | \bar{y}_v{}^{(h)}(a,v)|]dv,$$

Consider a typical one of these four terms; the other terms will satisfy similar inequalities:

$$\begin{aligned}\int_a^b | \bar{y}_u{}^h(u,d)| \,du &= (1-h)\int_{a(1-h)}^{b(1-h)} | y_u{}^{(h)}[u(1-h), v(1-h)] \,| \,du \\ &\leq \frac{1-h}{h^2}\int_{a(1-h)}^{b(1-h)}\int_{u(1-h)}^{u(1-h)+h}\int_{d(1-h)}^{d(1-h)+h} | y_\xi\,(\xi,\eta)| \,d\xi d\eta du \\ &\leq \frac{1-h}{h}\int_{d(1-h)}^{d(1-h)+h}\int_0^1 | y_u(u,\eta)| \,d\eta \\ &= \frac{1-h}{h}\int_{d(1-h)}^{d(1-h)+h} V_0{}^{(u)1}[y(u,\eta)]d\eta < 2V_0{}^{(u)1}[y(u,d)] < \infty,\end{aligned}$$

for h sufficiently small, $V_0{}^{(u)1}[y(u,V)]$ being metrically continuous at $V = d$.

THEOREM 1. *If S, $S : x^i = x^i(u,v)$, $i = 1, \cdots, N$, is a surface of class L,*

$$L(S) = G(S) = \int_0^1\int_0^1 \sqrt{EG - F^2}\,dudv.$$

Proof. Using the elementary inequality,

$$\left| \sqrt{\sum_{i=1}^{N}(\alpha_i)^2} - \sqrt{\sum_{i=1}^{N}(\beta_i)^2} \right| \leq \sum_{i=1}^{N} | \alpha_i - \beta_i |,$$

we see immediately that (using the fact that the representation is of class L)

$$\lim_{h\to 0}\iint_Q \left| \sqrt{EG-F^2} - \sqrt{\bar{E}_h\bar{G}_u - \bar{F}_h{}^2} \right| dudv = 0$$

and thus that

$$L(S) \leq \iint_Q \sqrt{EG-F^2}\, dudv,$$

since $\bar{S}_h \to S$, and

$$L(\bar{S}_h) = \iint_Q \sqrt{\bar{E}_h\bar{G}_u - \bar{F}_h{}^2}\, dudv,$$

$\bar{S}_h$ being of class C'.*

Since the $x^i(u, v)$ are all A. C. T., we can find a sequence, $\{T_n\}$, of subdivisions of Q, by means of lines parallel to the axes, into rectangles $R^{(n)}_{i,j}$ of diameter $< \delta_n$, where $\lim_{n\to\infty} \delta_n = 0$, so that each pair of functions $x^i(u, v)$ and $x^j(u, v)$ satisfy the hypotheses of Lemma 4 on the boundary of every $R^{(n)}_{i,j}$.† Then if $R^{(n)}_{i,j}$ is any rectangle and $F(R^{(n)}_{i,j})$ is the Geöcze set function defined in § 2, it is clear from Lemma 4 that (if $R^{(n)}_{i,j}$ is interior to Q)

$$F(R^{(n)}_{i,j}) \geq \sqrt{\sum_{k=2}^{N}\sum_{l=1}^{k-1}\left[\iint_{R^{(n)}_{i,j}} \frac{\partial(x^k, x^l)}{\partial(u, v)}\, dudv\right]^2}$$

and thus

$$G(S) \geq \sum_{i,j} \sqrt{\sum_{k=2}^{N}\sum_{l=1}^{k-1}\left|\iint_{R^{(n)}_{i,j}} \frac{\partial(x^k, x^l)}{\partial(u, v)}\, dudv\right]^2}. \tag{4.2}$$

But clearly the right side of (4.2) approaches $\iint_Q \sqrt{EG-F^2}\, dudv$ as n becomes infinite. Thus

$$L(S) \geq G(S) \geq \iint_Q \sqrt{EG-F^2}\, dudv,$$

and the theorem is completely demonstrated.

5. *Extensions to n-dimensions.* All of the preceding results together with their proofs (and no doubt those of the next two sections) can be generalized verbatim (essentially) to n-dimensional manifolds. We shall merely state the corresponding definitions:

---

* Lebesgue, *loc. cit.* (1), pp. 313-314. A surface of class C' is one for which all the representing functions are continuous together with their first partial derivatives.

† Which lies entirely interior to Q. In the rest of the proof, we need consider only such rectangles.

Definition 1. A function, $f(x^1, \cdots, x^n)$, is said to be *B. V. T.* in $(0, \cdots, 0; 1, \cdots, 1)$ if

(i) for almost all $(X^1, \cdots, X^{i-1}, X^{i+1}, \cdots, X^n)$, $f(X^1, \cdots, X^{i-1}, x^i, X^{i+1}, \cdots, X^n)$ is of bounded variation in x^i, $i = 1, \cdots, n$.

(ii) $V_0{}^{(x^i)1}[f(X^1, \cdots, X^{i-1}, x^i, X^{i+1}, \cdots, X^n)] < \mu(X^1, \cdots, X^{i-1}, X^{i+1}, \cdots, X^n)$, $\int_0^1 \cdots \int_0^1 \mu(s^1, \cdots, s^{n-1}) ds^1 \cdots ds^{n-1} < \infty$.

Definition 2. A function, $f(x^1, \cdots, x^n)$, is said to be *A. C. T.* in $(0, \cdots, 0; 1, \cdots, 1)$ if

(i) it is continuous and B. V. T.

(ii) for almost all $(X^1, \cdots, X^{i-1}, X^{i+1}, \cdots, X^n)$, $f(X^1, \cdots, X^{i-1}, x^i, X^{i+1}, \cdots, X^n)$ is absolutely continuous in x^i, $i = 1, \cdots, n$.

Definition 3. Given $f(x^1, \cdots, x^n)$, we define

$$f^{h^1, \cdots, h^n}(x^1, \cdots, x^n) = \frac{1}{h^1, \cdots, h^n} \int_{x^1}^{x^1+h^1} \cdots \int_{x^n}^{x^n+h^n} f(\xi^1, \cdots, \xi^n) d\xi^1 \cdots d\xi^n,$$

$$\bar{f}^h = f^{h, \cdots, h}[x^1(1-h), \cdots, x^n(1-h)].$$

Definition 4. We define the interval functions, $\alpha^i(I)$, replacing the $\alpha(I)$ and $\beta(I)$ of § 1, by

$$\alpha^i(I) = \int_{a^1}^{b^1} \cdots \int_{a^{i-1}}^{b^{i-1}} \int_{a^{i+1}}^{b^{i+1}} \cdots \int_{a^n}^{b^n} [f(x^1, \cdots, x^{i-1}, b^i, x^{i+1}, \cdots, x^n)$$
$$- f(x^1, \cdots, x^{i-1}, a^i, x^{i+1}, \cdots, x^n)]\, dx^1 \cdots dx^{i-1} dx^{i+1} \cdots dx^n,$$
$$(i = 1, \cdots, n),$$

$$I : (a^1, \cdots, a^n; b^1, \cdots, b^n) : a^i \leq x^i \leq b^i, \qquad (i = 1, \cdots, n).$$

Definition 5. We define the *x^i-variation* $V_{Q^{(n)}}^{(x^i)}(f)$ as the variation of the interval function $\alpha^i(I)$.

Definition 6. We define the *Fréchet distance* of two manifolds as suggested in § 2, and consider two manifolds, M_1 and M_2, equal if $\| M_1, M_2 \| = 0$, and say that the sequence of manifolds, $\{M_n\}$, approaches the manifold M if $\lim_{n\to\infty} \| M, M_n \| = 0$.

Definition 7. The definitions of an *n-polyhedron,* Π, and thus of the "*Lebesgue volume*" of an n-dimensional manifold as well as those of the *order of a point* * with respect to a manifold and thus of the "*Geöcze volume*" of an n-dimensional manifold are the precise analogs of the corresponding definitions for $n = 2$.

Definition 8. We say that an n-dimensional manifold, M, $M : x^i = x^i(u^1, \cdots, u^n)$, $i = 1, \cdots, N$, is of $(n+1)$*-dimensional measure zero* if every manifold, $M_{i_1, \ldots, i_{n+1}} : x^i = x^i(u^1, \cdots, u^n)$, $i = i_1, i_2, \cdots, i_{n+1}$, obtained from M as indicated is of $(n+1)$-dimensional measure zero.

Definition 9. A manifold M, $M : x^i = x^i(u^1, \cdots, u^n)$, $i = 1, \cdots, N$, is of *class L* if

(i) the $x^i(u^1, \cdots, u^n)$ are all A. C. T.;

(ii) $$\lim_{h \to 0} \int \cdots \int_{Q^{(n)}} \left| \frac{\partial(\bar{x}_h^{i_1}, \cdots, \bar{x}_h^{i_n})}{\partial(u^1, \cdots, u^n)} - \frac{\partial(x^{i_1}, \cdots, x^{i_n})}{\partial(u^1, \cdots, u^n)} \right| du^1 \cdots du^n =$$
$$i_1, \cdots, i_n = 1, \cdots, N;$$

(iii) for almost every U^j the manifold, $x^i = x^i(u^1, \cdots, u^{j-1}, U^j, u^{j+1}, \cdots, u^n)$, is of n-dimensional measure zero, $j = 1, \cdots, n$ (no doubt this last condition is a consequence of the first two but this is as yet unproved).

6. *Generalized conformal representations of surfaces.*

Definition. We say that the surface S, $S : x^i = x^i(u, v)$, $i = 1, \cdots, N$, $(u, v) \epsilon R$, is represented generalized conformally on a Jordan region, R, in the plane if

(1) the $x^i(u, v)$ are all A. C. T. in R, with $(x_u{}^i)^2$ and $(x_v{}^i)^2$ summable over the interior of R, $i = 1, \cdots, N$,

(2) $\mathrm{E} = \mathrm{G}$, $\mathrm{F} = 0$ almost everywhere interior to R.

THEOREM 1. *If S is represented generalized conformally on a Jordan region R,*

$L(S_h) \leq L(S)$, $S_h : x^i = x_h{}^i(u, v)$,

$$x_h{}^i(u, v) = \frac{1}{h^2} \int_u^{u+h} \int_v^{v+h} x^i(\xi, \eta) \, d\xi d\eta, \quad (u, v) \epsilon R_h, \qquad i = 1, \cdots, N,$$

* See for instance, Tannery, *Introduction à la theorie des fonctions d'une variable,* vol. 2, note by Hadamard.

where R_h is that Jordan region bounded by a simple closed rectifiable curve all of whose points are at a distance $\geq h\sqrt{2}$ from the boundary, C, of R and are joinable to a point P_0 of R by means of an arc all of whose points are at a distance $\geq h\sqrt{2}$ from *C, P_0 being a preassigned fixed (for all h) interior point of R.*

Proof. It is clear that the representation is of class L, so that

$$L(S) = \tfrac{1}{2} \iint\limits_R (E + G)\,dudv = \tfrac{1}{2} \sum_{i=1}^{N} \iint\limits_R [(x_u{}^i)^2 + (x_v{}^i)^2]\,dudv$$

Now

$$L(S_h) \leq \tfrac{1}{2} \iint\limits_{R_h} (E_h + G_h)\,dudv = \tfrac{1}{2} \sum_{i=1}^{N} \iint\limits_{R_h} \left[\left(\frac{\partial x_h{}^i}{\partial u}\right)^2 + \left(\frac{\partial x_h{}^i}{\partial v}\right)^2\right] dudv$$

FURTHER LITERATURE.

G. A. Maggi, "Sull' area del superficie curve," *Atti della Reale Accademia dei Lincei*, Ser. 5, Vol. 5 (1896), pp. 440-445.

C. Juel, "Om bestemmelsen og arealer og volumer," *Nyt Tidsskrift for Mathematik*, Vol. 8 (1897), pp. 49-59.

H. Lebesgue, "Sur la définition de l' aire d' une surface," *Comptes Rendus*, Vol. 129 (1899), pp. 870-873; "Sur la définition de certaines intégrales de surface, etc.," *ibid.*, Vol. 131 (1900), pp. 867-870, 935-937.

O. Stolz, "Zur Erklärung der Bogenlänge, etc.," *Transactions of the American Mathematical Society*, Vol. 3 (1902), pp. 23-37.

C. de la Vallée Poussin, "Sur la définition de l' aire, etc.," *Annales de la Société Scientifique de Bruxelles*, Vol. 27 (1902-3), pp. 90-91.

Z. de Geöcze, "Contributions à la quadrature des surfaces," *Comptes Rendus*, Vol. 152 (1911), pp. 678-679, Vol. 154 (1912), pp. 1211-1213; "Über die Quadratur der Flächen," *Mathematikai és Physikai Lapok*, Vol. 20 (1911), pp. 255-301, Vol. 21 (1912), pp. 25-57; "Zur Theorie der Quadratur von krummer Oberflächen," *Mathematikai és Termeszettudomanyi Ertesitö*, Vol. 31 (1913), pp. 306-318; "La quadrature des surfaces courbes," *Mathematische und Naturwissenschaftliche Berichte aus Ungarn*, Vol. 26 (1913), pp. 1-88; "Recherches générales sur la quadrature des surfaces courbes," *ibid.*, Vol. 27 (1914), pp. 1-21, 131-163, Vol. 30 (1916), pp. 1-29; "Über die allgemeine Fläche," *Mathematikai és Termeszettudomanyi Ertesitö*, Vol. 35 (1917), pp. 359-360.

K. Popoff, "Sur la notion de l' aire d' une surface," *Archiv für Mathematik und Physik*, Ser. 3, Vol. 26 (1917), pp. 18-23.

W. H. Young, "On a formula for an area," *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 18 (1919), pp. 339-374; "The triangulation method of defining the area of a surface," *ibid.*, Vol. 19 (1920), pp. 117-152; "On a new set of conditions for a formula for an area," *ibid.*, Vol. 21 (1922), pp. 75-94.

J. C. Burkill, "Expression of area as an integral," *ibid.*, Vol. 22 (1923), pp. 311-336.

G. Lampariello, "Sulle superficie continue che ammettono area finita," *Atti della Reale Accademia dei Lincei*, Ser. 6, Vol. 3 (1926), pp. 294-298.